

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 15, 2003, 05:00:47 ; Search time 4276 Seconds  
(without alignments)  
3239.695 Million cell updates/sec

Title: US-09-960-643-2  
Perfect score: 2513  
Sequence: 1 MGRKEDDCSSKKQTTNIR.....VKASGSSHCRAQTGVCLIM 476

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09960643/runat\_07032003\_090517\_19877/app\_query.fasta\_1.647  
-DB=GenEmbl -QFMT=fastp -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09960643.@C@N\_1\_1\_2496\_@runat\_07032003\_090517\_19877 -NCPU=6 -ICPU=3  
-NO\_XLPAY -NO\_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg\_hum.\*  
31: em.htg\_inv.\*  
32: em.htg\_other.\*  
33: em.htg\_mus.\*  
34: em.htg\_pln.\*  
35: em.htg\_rod.\*  
36: em.htg\_mam.\*  
37: em.htg\_vrt.\*  
38: em.sy.\*  
39: em.htgo\_hum.\*  
40: em.htgo\_mus.\*  
41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2513	100.0	1738	9 HS272L161	AL049688 Human gen
2	2513	100.0	2447	6 AX399682	AX399682 Sequence
3	2513	100.0	2474	9 BC032787	BC032787 Homo sapi
4	2509	99.8	2464	9 AF428261	AF428261 Homo sapi
5	2506	99.7	2612	9 AK095713	AK095713 Homo sapi
6	2327.5	92.6	2416	10 BC021840	BC021840 Mus muscu
7	2327.5	92.6	2427	10 AF428262	AF428262 Mus muscu
8	1612	64.1	1013	10 D86557	D86557 Rattus norv
9	1246	49.6	1074	6 AX166517	AX166517 Sequence
10	1246	49.6	1074	6 AX167587	AX167587 Sequence
11	1246	49.6	1158	6 AX167585	AX167585 Sequence
12	1246	49.6	1579	9 AF286366	AF286366 Homo sapi
13	1246	49.6	1671	6 AX167589	AX167589 Sequence
14	1157	46.0	1448	10 BC014825	BC014825 Mus muscu
15	1155.5	46.0	1439	10 RATCAMPKAA	L28288 Rattus norv
16	1154	45.9	1480	9 HUMCKI	L41816 Homo sapien
17	1151	45.8	1402	10 RATPRKI	L24907 Rattus norv
18	1070	42.6	1618	10 AB004267	AB004267 Rattus no
19	1068	42.5	1332	10 D86556	D86556 Rattus norv
20	1065	42.4	1554	10 AF181984	AF181984 Mus muscu
21	1064	42.3	1032	10 AB023027	AB023027 Mus muscu
22	1041	41.4	1032	6 AX166520	AX166520 Sequence
23	1004	40.0	1546	3 AB021864	AB021864 Caenorhab
24	990	39.4	1282	6 AR139102	AR139102 Sequence
25	957.5	38.1	3316	3 DME17917	Y17917 Drosophila
26	865	34.4	157875	9 HS272L16	AL023754 Human DNA
27	840	33.4	2202	9 BC033746	BC033746 Homo sapi
28	809.5	32.2	42404	3 CBRC46106	AC084656 Caenorhab
29	753	30.0	2897	6 AX099935	AX099935 Sequence
30	747.5	29.7	1503	6 AX239863	AX239863 Sequence
31	747.5	29.7	2297	6 AX239861	AX239861 Sequence
32	747.5	29.7	3003	9 BC001921	BC001921 Homo sapi
33	747.5	29.7	3012	9 BC000497	BC000497 Homo sapi
34	747.5	29.7	3012	9 BC005828	BC005828 Homo sapi
35	747.5	29.7	3018	9 BC017363	BC017363 Homo sapi
36	747.5	29.7	3057	9 BC019256	BC019256 Homo sapi
37	747.5	29.7	3124	6 AX350342	AX350342 Sequence
38	746	29.7	3141	10 BC017634	BC017634 Mus muscu
39	741	29.5	3655	10 RATCBVA	L22557 Rattus norv
40	735	29.2	182054	10 AL365314	AL365314 Mouse DNA
41	735	29.2	184754	2 AC022675	AC022675 Mus muscu
42	699.5	27.8	955	3 DDIMLCK	M64176 D.discoideu
43	697.5	27.8	3165	10 MMCDPK4	X58995 Mouse mRNA
44	692.5	27.6	1706	10 S65840	S65840 Ca2+/calmod
45	692.5	27.6	1785	10 RATCALM	M64757 Rat Ca++/ca

ALIGNMENTS

RESULT 1

```

HS272L161
LOCUS       HS272L161               1738 bp    mRNA    linear    PRI 21-APR-1999
DEFINITION   Human gene isolated from PAC 272L16, chromosome 1, similar to
              calcium/calmodulin dependent protein kinases.
ACCESSION   AL049688
VERSION     AL049688.1  GI:4678721
KEYWORDS    .
SOURCE      Homo sapiens.
            Homo sapiens.
ORGANISM    Mammalia; Eutheraia; Primates; Catarrhini; Homnidae; Homo.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheraia; Primates; Catarrhini; Homnidae; Homo.
            Rhodes/S.
REFERENCE   1 (bases 1 to 1738)
AUTHORS     Direct Submission
TITLE       Submitted (21-APR-1999) E-mail contact: humquery@sanger.ac.uk
JOURNAL     This sequence was generated from cDNA clones isolated using
COMMENT     This sequence from the bacterial clone 272L16 (AL023754) and EST data.
            The EST sequences listed match this sequence with an identity of at
            least 95% between the coordinates shown.
            Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1/ Partial, experimentally
            determined gene.
FEATURES             Location/Qualifiers
     source          dJ272L16.C1.1.
                     1..1738
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="lq32.1-32.3"
                     <1..1448
                     /codon_start=3
                     /product="hypothetical protein"
                     /protein_id="CAB41259.1"
                     /db_xref="GI:4678722"
                     /translation="STLEAMGRKEEDDCSSWKQTTNIRKTFIFMEVLGSGAFSEVFL
                     VKQRLTGKLFALKCIKSPAPRDSLENEIAVLKKIKHENIVTLEDIYESTHYLVLM
                     QLVSGELFDRLRGVTEKDALSVIQVLSAVKYLHENGIVHRDLXPENLLLTPE
                     ENSKIMTDFGUSKMEQNGIMSTAGTGGVVAPELQAQPSYKAVDCWSIGVIYILL
                     CGYPPEYESKLEFKIEGYEFESFPWDDISESAKDFICHLEKDPNERYTCEKA
                     LSHPTWGTALHRDIYFSVLIQIKNFPAKSKWQAFNAAVVHHMKLHMLHS PGV
                     RPEVNRPPQOASRTSPSPSEITITEAPVLDHSAVPALTPCQHRRPRTAPGGR
                     SLNCLVNGSLHTSSSLVPMHOGSLAAGPCGCCSSCLNIGSKGKSSYCSSEPTLLKKANK
                     KQNFSEVMVPVKASGSSHCRAGTGVCLIM"
                     <1..1107
                     /number=1
                     108..236
                     /number=2
                     237..311
                     /number=3
                     312..450
                     /number=4
                     451..574
                     /number=5
                     575..650
                     /number=6
                     651..763
                     /number=7
                     653..777
                     /notes="matches EST A1215131 from clone IMAGE:1925595"
     exon           764..842
                     /number=8
                     843..930
                     /number=9
                     868..1159
                     /notes="matches EST AA351937"
     exon           931..1355
                     /number=10
                     1356..1483
                     /number=11
                     join(1469..1550,1553..1723,1721..1738)
     misc_feature   /notes="matches EST R05661 from clone 29500"
                     1484..1736
     exon           /number=12

```

---

```

BASE COUNT  432 a  491 c  458 g  357 t
ORIGIN
Alignment Scores:
Pred. No.:      2,33e-170      Length:      1738
Score:          2513.00        Matches:      476
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:    100.00%       Indels:         0
DB:              9            Gaps:          0

US-09-960-643-2 (1-476) x HS272L161 (1-1738)
Qy  1  MetGlyArgLysGluGluAAspCysSerSerTrpLysLysGlnThrThrAsnIleArg  20
    |||||
Db  18  ATGGTGCAGAAAGGAAGAGTGCAGTTCCTGGAAGAAACAGACCACCAACATCCGG  77
Qy  21  LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal  40
    |||||
Db  78  AAAACCTTCATTTTATGGAAGTCTGGGATCGTGGAGCTTTCTCAGAAAGTTTCTCGTGGTG  137
Qy  41  LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe  60
    |||||
Db  138  AAGCAAGACGTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCAACCTGCCTTC  197
Qy  61  ArgAspSerSerLeuGluAAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle  80
    |||||
Db  198  CGGACAGCAGCCTGGAGATGAGATTGCTGTGTGAAAAAGATCAAGCATGAAACATT  257
Qy  81  ValThrLeuGluAAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal  100
    |||||
Db  258  GTGACCTCGGAGGACATCTATGAGAGCACCACCCACTACTACCTGGTCATGCAGCTTGT  317
Qy  101  SerGlyGlyGluLeuPheAAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla  120
    |||||
Db  318  TCTGTGGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAGAAGGATGCC  377
Qy  121  SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAAsnGlyIleVal  140
    |||||
Db  378  AGTCTGGTGATCCAGCAGCTCTTGTGGCAGTGAATACCTACATGAGATGGCATCGTC  437
Qy  141  HisArgAspLeuLysProGluAAsnLeuLeuTyrLeuThrProGluGluAAsnSerLysIle  160
    |||||
Db  438  CACAGAGACTTAAGCCCGAAGAACTGCTTTACCTTACCTTGAAGAGACTCTAAGATC  497
Qy  161  MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys  180
    |||||
Db  498  ATGATCACTGACTTTGGTCTGTCCAAGATGGAACAGAAATGGCATCATGTCCACTGCCTGT  557
Qy  181  GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal  200
    |||||
Db  558  GGGACCCCGAGGCTAGTGGCTCCAGAGAGTGTGGCCCAAGAAACCTACAGCAAGGCTGTG  617
Qy  201  AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyr  220
    |||||
Db  618  GATTGCTGCTCCATCGGCGTCATCACCTACATATTGCTCTGTGGATACCCCCCATTCAT  677
Qy  221  GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer  240
    |||||
Db  678  GAAGAAACCGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAGTTTGGAGTCT  737
Qy  241  ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys  260
    |||||
Db  738  CCATTCTGGGATGACATTTCTGAGTCAGCCCAAGGACTTTATTGGCCACTTCTGTTGAGAAG  797
Qy  261  AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn  280
    |||||
Db  798  GATCCGAACGAGCGGTACACCTGTGAGAAGGCTTGTAGTCATCCTCGGATTCGACGGAAC  857
Qy  281  ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla  300
    |||||
Db  858  ACAGCCCTCCACCGGAGACATCTACCCATCAGTCAGCTCCAGCTCCAGAGAACTTTGCT  917

```

```
Qy 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArgLysLeu 320
Dy 918 AAGAGCAAGTGGAGGCAAGCTTCAACGAGCAGAGCTGTGGTGCACCACATGAGGAAGCTA 977
Qy 321 HisMetAsnLeuHisSerProGlyValArgProGluValValGluAsnArgProGluThr 340
Dy 978 CACATGAACCTGGACAGCCGGCGCTCCGCCAGAGGTGGAGAACAGAGCCGCTGAAACT 1037
Qy 341 GlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaProVal 360
Dy 1038 CAACCTCAGAACCTCTAGACCAGCTCCCTGGAGATCACCATCACCGAGGACCTGTC 1097
Qy 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380
Dy 1098 CTGACACACAGTGTAGCAGCTCCCTGGCCCTGACCAATTACCTGCGCAGCATGGCCGCCGG 1157
Qy 381 ProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400
Dy 1158 CCCACTGCCCTGTGGCAGTCCCTCACTGGCTGGTCAATGGCTCCCTCCACATCAGC 1217
Qy 401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420
Dy 1218 AGCAGCTGTGTGCCATGCATCAGGGTCCCTGGCCGCCGGCCCTGTGGCTGCTCC 1277
Qy 421 SerCysLeuAsnIleGlySerLysGlyLysSerSerSerSerSerSerSerSerSerSer 440
Dy 1278 AGCTGCCCTGAACATTTGGAGCAAGGAAGTCTCTCTACTGCTGCTGAGCCACACTCTCTC 1337
Qy 441 LysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAlaSer 460
Dy 1338 AAAAGGCCAACAAAAACAGAACTTCAAGTCGAGGTCATGGTACCAGTTAAAGCCAGT 1397
Qy 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476
Dy 1398 GGCAGCTCCCACTGCCGGGAGGCGACACTGGAGTGTCTCTATTATG 1445
RESULT 2
AX399682 2447 bp DNA linear PAT 06-JUN-2002
LOCUS
DEFINITION Sequence 3 from Patent WO0224947.
ACCESSION AX399682
VERSION AX399682.1 GI:21335455
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Delaney,A.D. and Yoganathan,T.
Cancer associated protein kinases and their uses
Patent: WO 0224947-A 3 28-MAR-2002;
KINETEK PHARMACEUTICALS INC (CA); UNIV BRITISH COLUMBIA (CA)
FEATURES
source
1. .2447
/organism="Homo sapiens"
/db_xref="taxon:9606"
70. .1500
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD33521.1"
/db_xref="GI:21335456"
translation="MGRKEEDDCSSKKQTTNIRKTFIMFVLGSGAFSEVFLVKQRL
TGKLFALKCIKSPAFDSLENIATLKKIHENIVLEDIYESTHYIYLMQLVSG
GELFDILRGVYTEKDSALVIQQVL SAVKY LHENGIVHRDLKPNLLYLPENSKI
MITDFGLSKMEONGINSTAGTGPVVAPEVLAQKPYSKAVDCWSITGVITYILCGYPP
FYETESKLPEKIKEGYVEFESPFWDLLSESAKDFICHLEKDPNERYTCEKALSHPW
IDGNTALHRDIPSLSLQIKNFPAKSRQAFNAAYVHHMRKLHNNLHSGYRPEVE
NRPEFOAETSPESPEITTEAPVLDHVALPALQTLPQOHRGRTPAPGRSUNCL
VNGSLHISLSPMHQGSLLAAGPCGCCSCLNIGSKGSSYCSSEPTLLKANKKQNEK
SEYMPVPAKSGSHCRAGQTGVCLIM"
BASE COUNT 590 a 707 c 604 g 546 t
ORIGIN
```

```
Alignment Scores: 3.62e-170 Length: 2447
Pred. No.: 2513.00 Matches: 476
Score: 2513.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 6 Gaps: 0
DB:
US-09-960-643-2 (1-476) x AX399682 (1-2447)
Qy 1 MetGlyArgLysGluGluAspAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
Dy 70 ATGGGTGGAAGGAGGAGATGACTGCAGTTCCTGGAAGAAACAGACCACCAATCCCG 129
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Dy 130 AAAACCTTCATTTTATGTAAGTCTGGGATCAGGAGCTTCTCAGAAGTTTTCCTGGTG 189
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
Dy 190 AAGCAAAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAGTCACTGCCTTC 249
Qy 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
Dy 250 CGGACACAGCCTGGAGATGAGATTGCTGTGTGAAAGATCAAGCATGAACATTT 309
Qy 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
Dy 310 GTGACCTTGGAGGACATCTATGAGACGACCCACCTACTACTGCTGTCATGCAGCTTGT 369
Qy 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
Dy 370 TCTGGTGGGAGCTCTTTGACCGATCCTGGAGCGGGTGTCTACACAGAGAGGATGCC 429
Qy 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
Dy 430 AGTCTGTGTATCCAGCAGCTTGTGCGCAGTGAATACCTATACATGAGAAATGTCATCGTC 489
Qy 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
Dy 490 CACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCTGAAGAACTCTAGATC 549
Qy 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
Dy 550 ATGATCATCTGACTTTGCTGTCTCCAGATGAACAGAAATGGCATCATGTCCTGCTGT 609
Qy 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Dy 610 GGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCAGAAACCCCTACAGAGGCTGTG 669
Qy 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProProPheTyr 220
Dy 670 GATTGCTGGTCCATCGGGCTCATCCTACATATATTGCTCTGTGATACCCCGCTTCTAT 729
Qy 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
Dy 730 GAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTACTATGATTTGAGTCT 789
Qy 241 ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
Dy 790 CCATCTGGGATGACATTTCTGAGTCAGCCCAAGACCTTATTTGCCACTTTCCTTGAGAAG 849
Qy 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280
Dy 850 GATCCGAACGAGCGGTACACCTGTGAGAAGCCCTTGAGTCATCCCTGGATGACGGAAC 909
Qy 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300
Dy 910 ACGGCCCTCCACCGGGACATCTACCCATCAGCTCAGCTCCAGATCCAGAGAACTTTGCT 969
Qy 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArgLysLeu 320
Dy 970 AAGACCAAGTGGAGCAAGCCTTCAACGCAGCAGCTGTGGTGCACCACATGAGGAAGCTA 1029
```

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 68 Row: k Column: 17  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 14196444.

FEATURES  
source  
location/Qualifiers  
1. .2474  
/organism="Homo sapiens"  
/db\_xref="LocusID:57172"  
/db\_xref="taxon:9606"  
/clone="MGC:44894 IMAGE:5179957"  
/tissue\_type="Brain, Lung, Testis, adult, pooled whole"  
/clone\_lib="NIH\_MGC\_115"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
70. .1500  
/codon\_start=1  
/product="calmodulin-dependent protein kinase Ig"  
/protein\_id="AAH32787.1"  
/db\_xref="GI:21619665"  
TGTLFALKCIKKSPAFRSDSLENEIAVLKKIKHENIVTLVDIYESTHYVLMQLVSG  
MINTDFGLSKMEQNGIMSTAGCTPGYVAPEVLAOKPYSKAVDCWSIGVITVILCGYPP  
FYBETESKLEFEKIEGYEYESPFWDIISAKDFICHLEKDPNERVTCERKALSHPM  
IDGNTALURDIIYFSVLQIQNFAKSKWRQFNAAAVVHHMRKLHNNLSHPGVPEVE  
NRPPETQASETSPPSEITITEAPVLDHSEVALPALTQPCQHGRRPTAPGGHSLNCL  
VNGSLHISSSLVPMHQGSLAAGCCSCNLNIGSKGKSYCSYCEPTLLKANKKQNFK  
SEVMVPVKASGSSHCRAGQTGVCILIM"

BASE COUNT 618 a 707 c 603 g 546 t

ORIGIN

Alignment Scores:  
Pred. No.: 3.67e-170 Length: 2474  
Score: 2513.00 Matches: 476  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-960-643-2 (1-476) x BC032787 (1-2474)

QY 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20  
Db 70 ATGGGTCCGAAGAAGAGATGACTGCTGCTGGAAGAAGAACAGACCAACATCCGG 129  
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
Db 130 AAAACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTTCTCAGAAGTTTCTCGTG 189  
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60  
Db 190 AAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCAATCAAGAAGTCACTGCCTTC 249  
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80  
Db 250 CGGGACACAGCTGGAGAAATGAGATGCTGTTGAAAAAGATCAAGCATGAACATT 309  
QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100  
Db 310 GTGACCTTGGAGGACATCTATGAGAGACCAACCCACTACTACCTGGTCATGCAGCTTGT 369  
QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120  
Db 370 TCTGTTGGGGAGCTCTTTGACCGGATCTCGAGCGGGGTGCTACACAGAGAAGATGCC 429  
QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140  
Db 430 AGTCTGGTGATCCAGCAGCTCTTGTGCGCAGTGAATACCTACATGAGAATGGCATCGTC 489  
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160  
Db 490 CACAGAGACTTAAAGCCCGAAACCTGCTTTTACCTTACCTTACCTTACCTTACCTTACCTT 549

321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProProGluThr 340  
|||||  
1030 CACATGAACCTGCACACCCGGCGTCCGCCAGAGGTGGAGACAGCGCGCTGAACCT 1089  
|||||  
341 GlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaProVal 360  
|||||  
1090 CAAGCCTTCAGAAACCTTAGACCCAGCTCCCTCGAGATCACCATCACCAGGACCGCTGTC 1149  
|||||  
361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380  
|||||  
1150 CTGACCACAGTAGACATCTCCCTGCGCTGACCAATACCTTCGCGAGATGGCGCGCG 1209  
|||||  
381 ProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400  
|||||  
1210 CCCACTGCCCTGGTGCAGCTCCCTCAACTGCTGCTCAATGGCTCCCTCCACATCAGC 1269  
|||||  
401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420  
|||||  
1270 AGCAGCTGTGCTGCATGATCAGGGTCCCTGGCGCGCGCGCTGTGCTGCTGCTCC 1329  
|||||  
421 SerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCysSerGluProThrLeuLeu 440  
|||||  
1330 AGTGGCTGAACATTGGGACCAAGAAAGTCTCTACTGCTCTGAGCCACACTCCTC 1389  
|||||  
441 LysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAlaSer 460  
|||||  
1390 AAAAAAGCCCAACAAAAACAGAACTTCAAGTCGGAGGTGATGTTACCACTTAAAGCCAGT 1449  
|||||  
461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476  
|||||  
1450 GGCAGCTCCCACTCGCGGCGAGGACACTGGAGTCTGTCTCATTTATG 1497  
|||||

RESULT 3  
BC032787  
LOCUS  
DEFINITION Homo sapiens, calcium/calmodulin-dependent protein kinase IG, clone  
MGC:44894 IMAGE:5179957, mRNA, complete cds.  
ACCESSION  
VERSION BC032787.1 GI:21619664  
KEYWORDS MGC.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2474)  
Direct Submission  
Submitted (06-JUN-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
Akhtar, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

```
Qy 161 MetIleThrAspPheGlyLeuSerLysMetGluAlaAsnGlyIleMetSerThrAlaCys 180
Db 550 ATGATCACTAGCTTTGGTCTGTCTCAAGATGGAACAAGATGGCATCATGTGCCACTGCTGT 609
Qy 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Db 610 GGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCAGAAACCCCTACAGCAAGGCTGTG 669
Qy 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProProPheTyr 220
Db 670 GATTGCTGGTCCATCGCGCTCATCACCTACATATTTGCTGTGTGATACCCCGGTCTAT 729
Qy 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
Db 730 GAAAGAACGGAGTCTAAGCTTTTCGAGAGATCAAGAGGGCTTACTATGAGTTGAGTCT 789
Qy 241 ProPheTrpAspPheSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
Db 790 CCATCTGGATGATACATTTCTGAGTCAGCCAGGACTTTATTGGCCACTTGTCTGAGAAG 849
Qy 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280
Db 850 GATCCGAACGAGCGGTACACTGTGAGAGGCTTGTGATCATCCCTGGATTGACGGAAC 909
Qy 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300
Db 910 ACAGCCCTCCACCGGGACATCTACCCATCATAGTCAGCTCCAGATCCAGAGAACTTTGCT 969
Qy 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArgLysLeu 320
Db 970 AAGAGCAAGTGAGGCAAGCTTCAACGCCAGCAGCTGTGGTGCACCACATGAGGAAGCTA 1029
Qy 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProProGluThr 340
Db 1030 CACATGAACCTGCGACAGCCCGGGCTCCGCCAGAGGTGGAGAACAGGCGCGCTGAACT 1089
Qy 341 GlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaProVal 360
Db 1090 CAAGCCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCACCATCACCAGGACACCTGTC 1149
Qy 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380
Db 1150 CTGACCAACAGTGTAGCAGCTCCCTGGCCCTGACCCAAATTAACCTGCCAGCATGCCGCCG 1209
Qy 381 ProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400
Db 1210 CCCACTGCCCTGTGGCAGGTCCCTCACTGCTGGTCAATGCTCCCTCCACATCAGC 1269
Qy 401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420
Db 1270 AGCAGCTGTGTGCCATGCATCAGGGTCCCTGCGCGCGGGCCCTGTGCTGCTGCTCC 1329
Qy 421 SerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCysSerGluProThrLeuLeu 440
Db 1330 AGCTGCTGACATTTGGAGCAAGGAAGTCTCTCTACTGCTCTGAGCCCACTCCCTC 1389
Qy 441 LysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAlaSer 460
Db 1390 AAAAAGGCCAACAAAAACAGACTTCAAGTCGGAGGTGATGTTACAGTTAAAGCCACT 1449
Qy 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476
Db 1450 GGCAGCTCCCACTCCGGGAGGGCAGACTGGAGTCTGTCTCATATTATG 1497

RESULT 4
AF428261
LOCUS AF428261 2464 bp mRNA linear PRI 07-NOV-2001
DEFINITION Homo sapiens calcium/calmodulin-dependent protein kinase I gamma
(CAMK1G) mRNA, complete cds.
ACCESSION AF428261
VERSION AF428261.1 GI:16755791
KEYWORDS
SOURCE Homo sapiens.
```

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2464)  
Schutte,B.C., Bjork,B.C., Coppage,K.B., Malik,M.I., Gregory,S.G.,  
Scott,D.J., Brentzell,L.M., Watanabe,Y., Dixon,M.J. and Murray,J.C.  
A preliminary gene map for the Van der Woude syndrome critical  
region derived from 900 kb of genomic sequence at 1q32-q41  
Genome Res. 10 (1), 81-94 (2000)  
20113118  
MEDLINE  
PUBMED 10645953  
REFERENCE 2 (bases 1 to 2464)  
Bjork,B.C., Watanabe,Y., Murray,J.C. and Schutte,B.C.  
Characterization of the human ortholog of rat Cam Kinase I gamma  
(CamK1g) at 1q32-q41  
Unpublished  
JOURNAL 3 (bases 1 to 2464)  
AUTHORS Bjork,B.C., Watanabe,Y., Murray,J.C. and Schutte,B.C.  
TITLE Direct Submission  
JOURNAL Submitted (05-OCT-2001) Genetics, University of Iowa, 140 EMBR,  
Iowa City, IA 52242, USA  
FEATURES  
Location/Qualifiers  
1..2464  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="1q32-q41"  
1..2464  
/gene="CAMK1G"  
73..1503  
/gene="CAMK1G"  
/codon\_start=1  
/product="calcium/calmodulin-dependent protein kinase I  
gamma"  
/protein\_id="AAL28100.1"  
/db\_xref="GI:16755792"  
TGKFLKLCIKKSPAFRDSLEIAVLKKIKHENITVLTLEDIVESTHYLYVMQLVSG  
GEFLDRLERGVYTKDASLIVQVL SAVKYLHENGIVHRLDLKPNLLVLTPEENSKI  
MITDFGLSKMEONGIMSTAGCTGYVAPEVLAQPKYSKAVDCWSIGVIYIILCGYPP  
FYDETSEKLFKEIGYEYFDDIDSESADFICHLLEKDPNERYTCERALSHPW  
IDNTALHRDIYPSVLOKFNFAKSKRQAFNAAVVHMRKLMHNLHSPGIRPEVE  
NRPETQASETSPPSEITITPAVDHVSVALPTQLPCQHRGRRTPAGGRSLNCL  
VNGSLHISSSLVPMHQSGLAAGCGCCSCLNIGNKGSKSYCSEPTLLKANKKQNFK  
SEVMVPVKASGSSSHCRAGTGVCLIN"  
2415..2420  
polyA\_signal  
2415..2420  
/gene="CAMK1G"  
BASE COUNT 609 a 708 c 602 g 545 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 7,05e-170 Length: 2464  
Score: 2509.00 Matches: 474  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 99.58% Mismatches: 0  
Query Match: 99.84% Indels: 0  
DB: 9 Gaps: 0  
US-09-960-643-2 (1-476) x AF428261 (1-2464)  
Qy 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrAsnIleArg 20  
Db 73 ATGGGTGAAAGGAGAGATGACTGCGAGTTCCTGGAGAAACACAGACCATCCG 132  
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
Db 133 AAAACCTTCATTTTATGGAAGTCTGGATCAGGAGCTTCTCAGAAAGTTTCTCGGTG 192  
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60  
Db 193 AAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTCATCAAGAGTACCTGCCTC 252  
Qy 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80

```
Db 253 CGGACAGCAGCCCTGGAGAATGAGATTGCTGTGTGAAAAAGATCAAGCATGAAAAACATT 312
Qy 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
Db 313 GTGACCTGGAGGACATCTATGAGAGCACCACCCATCTACTACCTGGCTGATGCAGCTTGT 372
Qy 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
Db 373 TCTGGTGGGAGCCTTTGACCGGATCTCGAGCGGGTGTCTACACAGAGAAGGATGCC 432
Qy 121 SerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
Db 433 AGTCTGTGTATCCAGCAGGCTTGTTCGGCAGTGAATATACCTACATGAGAATGCGATCGTC 492
Qy 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
Db 493 CACAGAGACTTAAAGCCGAAACCTGCTTACCTTACCCCTGAAGAGAACTCTAAGATC 552
Qy 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
Db 553 ATGATCACTGACTTGTGCTGTCCAGATGGAACAGAAATGGCATCATGCTCCACTGCCTGT 612
Qy 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Db 613 GGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCGAGAAACCCCTACAGCAAGGCTGTG 672
Qy 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLysCysGlyTyrProProPheTyr 220
Db 673 GATTGCTGGTCCATCGGGCGATCACTACCTACATATGCTCTGTGGATACCCCGCTCTAT 732
Qy 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
Db 733 GAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAGTTTGAGTCT 792
Qy 241 ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
Db 793 CCATTCTGGGATGACATTTCTGAGTCAGCAAGGACTTTATTGGCCACTGTGCTTGAGAAG 852
Qy 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280
Db 853 GATCCGAACGAGCGGTACACTGTGAGAAGCGCTTGAGTCATCCCTGGATTGACGGAAC 912
Qy 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300
Db 913 ACAGCCCTCCACCGGGACATCTACCCATCAGTCAGCTCCAGATCCAGATCCAGAAACTTTGT 972
Qy 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu 320
Db 973 AAGACCAAGTGGAGGCAAGCCTTCAAGCAGCAGCGCGTGTGCACCATGAGGAAGCTA 1032
Qy 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProGluThr 340
Db 1033 CACATGAACCTGCACAGCCGGGATCCGCCAGAGTGAGAGTGGAGAACAGCGCCCTGAAACT 1092
Qy 341 GlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaProVal 360
Db 1093 CAACCCCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCACCATCAGGAGGACCTGTGC 1152
Qy 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380
Db 1153 CTGGACCACAGTGTAGCACTCCCTGCCCTGACCCAAATACCTGCCAGCATGCGCCGCG 1212
Qy 381 ProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400
Db 1213 CCCACTGCCCTGTGTGAGAGTCCCTCAACTGCTGCTCAATGGCTCCCTCCACATCAGC 1272
Qy 401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420
Db 1273 AGCAGCTGTGTGCCATGATCAGGGGTCTCTGGCCCGCGGCTGTGGCTGCTGCTCC 1332
Qy 421 SerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCysSerGluProThrLeuLeu 440
```

```
Db 1333 AGCTCCCTGAACATTTGGGAACAAGGAAAGTCTCTCTACTGCTGTGAGCCACACTCCTC 1392
Qy 441 LysLysAlaAsnLysLysGlnAsnPhelySerGluValMetValProValLysAlaSer 460
Db 1393 AAAAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTCAATGGTACCAAGTTAAAGCCAGT 1452
Qy 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476
Db 1453 GGCACGCTCCCACTCGCGGCGAGGCAGACTGGAGTCTGCTCTCATTTATG 1500
```

## RESULT 5

AK095713

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AUTHORS

JOURNAL

COMMENT

FEATURES

BASE COUNT

ORIGIN

Alignment Scores:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-960-643-2 (1-476) x AK095713 (1-2612)

Qy 1 MetGlyArgLysGluLysAspAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20

Db 246 ATGGGTCGAAAGGAAGATGACTGCATGCTCTCTGGAAGAAACAGACCACCAATCCGG 305

```
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
|||
Db 306 AAAACCTTCATTTTATGGAAGTGTGGATCAGGAGCTTTCTCAGAAAGTTTTCCTGGTG 365
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
|||
Db 366 AGCAAAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAGTCACTGCCTTC 425
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysIleLysHisGluAsnIle 80
|||
Db 426 CGGGACAGCAGCCTGGAGATCAGATTGCTGTGTGAAAGATCAAGCATGAACAAT 485
QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
|||
Db 486 GTGACCTCGAGGAGATCATATGAGAGCACCCACTACTACCTGGTTCATGCAGCTTGT 545
QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
|||
Db 546 TCTGTGGGGAGCTCTTTGACCGGATCTTGGAGCGGGGTGTCTACACAGAAAGATGCC 605
QY 121 SerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
|||
Db 606 AGTCTGTGTATCAGCAGCTCTGTGCGCAGTGAATACTACATGAGATGGCATCGTC 665
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
|||
Db 666 CACAGAGACTTAAGCCCGAAACCGCTTTACCTTACCCCTGAAGAACTCTAAGATC 725
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
|||
Db 726 ATGATCAGTACTTGGTCTGTCCAAAGATGGAACAGATGGCATCATGTCCACTGCCGT 785
QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
|||
Db 786 GGSACCCAGGCTAGTGGCTCCAGAAGTGTGGCCCAAGAACCTACAGCAAGCGTGTG 845
QY 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyr 220
|||
Db 846 GATTGCTGTCTCATCGGCTCATCACTACATATTATTTCTGTGGATACCCCTTCTAT 905
QY 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
|||
Db 906 GAAGAAACCGAGTCTAAGCTTTTCAGAGAATCAAGAGGCGCTACTATGAGTTGAGTCT 965
QY 241 ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
|||
Db 966 CCATTTCTGGGATGACATTTCTGAGTCAGCCCAAGGACTTTATTGCGCACTTCTTGAGAAG 1025
QY 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTyrIleAspGlyAsn 280
|||
Db 1026 GATCCGAACGAGCGGTACACTGTGAGAAGCGCTTGAGTCTATCCCTGGATTGACGGAAAC 1085
QY 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300
|||
Db 1086 ACAGCCCTCACCGGGACATCTACCATCAGTCAGCTCAGCCCTCCAGATCCAGAAACTTGTCT 1145
QY 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu 320
|||
Db 1146 AAGACAAGTGGAGGCAAGCCCTTCAACGAGCAGCTGTGGTGGCCACCATGAGGAAGCTA 1205
QY 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProProGluThr 340
|||
Db 1206 CACATGAACCTGCACAGCCCGCGCTCCGCCACAGAGGTGGAGACAGSCCCCTGAACT 1265
QY 341 GlnAlaSerGluThrSerArgProSerProGluIleThrIleThrGluAlaProVal 360
|||
Db 1266 CAAGCCTCAGAAACCTCTAGACCCAGCTCCCTCGAGATCACCATCACCGAGCACCTGTC 1325
QY 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArg 380
|||
Db 1326 CTGGACACAGTGTAGCACTCCCTGCCCTGACCCCAATTACCTCGCCAGCATGGCGCGCGG 1385
```

```
QY 381 ProThrAlaProGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400
|||
Db 1386 CCCACTGCCCTGGTGGCAGGTCCTCAACTGCCTGGTCAATGGCTCCTCCACATCAGC 1445
QY 401 SerSerLeuValProMethHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420
|||
Db 1446 AGCAGCTGGTGGCCCATGTCATCAGGGTCCCTGGCGCGCGGCTGTGGCTGCTGCCTCC 1505
QY 421 SerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCysSerGluProThrLeuLeu 440
|||
Db 1506 AGTGGCTGAACATTGGGAGCAAGAAAGTCTCTCTACTGCTGAGCCACACTCTCTC 1565
QY 441 LysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAlaSer 460
|||
Db 1566 AAAAGGCCCAACAAAAACAGAACTTCAAGTCGGAGGTTCATGGTACCAAGTTAAAGCCAGT 1625
QY 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476
|||
Db 1626 GGCAGCTCCCATYTCGCCGCGCAGGCAGACTGGAGTCTGTCTCATTTATG 1673
RESULT 6
LOCUS BC021840 2416 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, clone MGC:30513 IMAGE:4502479, mRNA, complete cds.
ACCESSION BC021840
VERSION BC021840.1 GI:18256866
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1. (bases 1 to 2416)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (18-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 41 Row: i Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: GenomeScan gene
prediction.
Location/Qualifiers
1. .2416
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="MGC:30513 IMAGE:4502479"
/tissue_type="Eye, retina, mouse strain C57BL\6"
/clone_lib="NTH_MGC_94"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
87. .1520
/codon_start=1
/product="Unknown (protein for MGC:30513)"
/protein_id="AAH21840.1"
/db_xref="GI:18256867"
CDS
```

/translation="MGRKEBEDSCSWKQNTNIRKTFIFMEVLGSGAFSEVFLVKORV  
TKGLFALCKIKKSAFRRDSSLENEIAVLKRKHENIVITLEDIYTHHLYLVMLVSG  
GELFDRLERGVYTKEDASLYIQOVLAVKYLHENGIVHRDLAPENLLYLPENSKI  
MITDFGLSKQNGVMGTACGTGYVAPELVAQKPYSAVDKWSIGVITYILLCGYPP  
FYETESKLFKIKEGYIEFEPFWDIISAKDFICHLLEKDPNERYTEKALRHPW  
IDGNTALHRDITVPSVLIQIKNFPAKSKWRAFNAAVVHMRKLMHNLHSPYRQVE  
NRPVSPAPEVSRDSDSSITEAPILDPSTPLPALPCLPSHSSRPSAPSGRSLNC  
LYNGLSLVPSKQNGPLATGCGCCSCLNIGNKSSYCSSEPTFLFRKANKQNF  
KSEWNPVKAAGSTHCRGGGTGVLVM"  
BASE COUNT 572 a 694 c 602 g 548 t  
ORIGIN

## Alignment Scores:

Pred. No.: 6, 24e-157 Length: 2416  
Score: 2327.50 Matches: 439  
Percent Similarity: 94.97% Conservative: 14  
Best Local Similarity: 92.03% Mismatches: 23  
Query Match: 92.62% Indels: 1  
DB: 10 Gaps: 1

US-09-960-643-2 (1-476) x BC021840 (1-2416)

QY 1 MetGlyArgGluGluAspCysSerSerTrpLysGlnThrThrAsnIleArg 20  
Db 87 ATGGGCGTAAGGAGGAGGACTGCAGTTCTCGAAGAAACAGACCACCAATCAGG 146  
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
Db 147 AMAACCTTCATCTTCATGGAAGTGTGGGATCAGAGCTTCTCAGAGGTGTTCCTGGTG 206  
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60  
Db 207 AAGCAAGAGTGACTGGGAACCTTTGCTCTGAAATGTATCAAGAACTCACCAGCCCTC 266  
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80  
Db 267 CGGACAGCAGCTAGAGAAATGAGATCGCTGTGTGAAGAGATCAAGCATGAGAACATT 326  
QY 81 ValThrLeuGluAspIleTrpGluSerThrHisTrpHisTrpValLeuValMetGlnLeuVal 100  
Db 327 GTGACCTGGAGGACATCTATGAGAGCACCACCCACTACTACCTGGTCATGACGCTTGT 386  
QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTrpThrGluLysAspAla 120  
Db 387 TCTGGAGGTGAGCTTTTGACCGGATCTTAGAGCTGTGTGTACACAGAAAGATGCC 446  
QY 121 SerLeuValIleGlnValLeuSerAlaValLysTrpLeuHisGluAsnGlyIleVal 140  
Db 447 AGCTGTGTCATCCAGAGCTGTGTCTGGGTGAATACCTTCATGAGATGGCATCGTC 506  
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTrpLeuThrProGluLysSerLysIle 160  
Db 507 CACAGAGATCTAAAGCCTGAAACCTGTGTACCTCAGCCCTCAGAGAGATCCCAAGATC 566  
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180  
Db 567 ATGATCAGCTGACTTTGGTCTATCCAGATGGAGAGATGAGTATGATCCACAGCTTGT 626  
QY 181 GlyThrProGlyTrpValAlaProGluValLeuAlaGlnLysProTrpSerLysAlaVal 200  
Db 627 GGGACCCAGGCTACGTGGCTCCAGAGTGTGGCCAGAGCCCTACAGTAAGGCTGTG 686  
QY 201 AspCysTrpSerIleGlyValIleThrTrpIleLeuLeuLysGlyTrpProPheTrp 220  
Db 687 GACTGTGTGCTGCTATCATACATACATGCTGTGTGGTATCCCCCTTTCAT 746  
QY 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTrpTrpGluPheGluSer 240  
Db 747 GAGAAACAGATCAAGCTTTTGGAGAGATCAAGAGGTTTACTACAGTTTTCAGTCT 806  
QY 241 ProPheTrpAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260  
Db 807 CCCTTCTGGGATGACATTTCTGAGTCAGCAAGGATTTTATTGTGCCATCTGTGGAGAAG 866

QY 261 AspProAsnGluArgTrpThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280  
Db 867 GACCCAAAGGACGGTACCTCGGAAAGCCCTCAGACACCCCTGGATTGATGAAC 926  
QY 281 ThrAlaLeuHisArgAspIleTrpSerValSerLeuGlnIleGlnLysAsnPheAla 300  
Db 927 ACAGCCCTGACCGGACATCTATCCATCTGTGAGCTCCAGATCCAGAAGAACTTTGCC 986  
QY 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArgLysLeu 320  
Db 987 AAGAGCAAGTGGAGCAAGCCTTCAACGACGAGCTGCGTGTGTGCATCAGAGAACTG 1046  
QY 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProProGluThr 340  
Db 1047 CACATGAACCTCAGACGCCCTGTCGCGCAAGAGGTGGAGAACAGGCCACTGTGTCC 1106  
QY 341 GlnAlaSerGluThrSerArgProSerProGluIleThrIleThrGluAlaProVal 360  
Db 1107 CTTGCTCTCAAGTCTCCAGACCCAGACTCCCATGACAGCTCCATCAGAGGCCCCCATC 1166  
QY 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380  
Db 1167 CTGACCCCAAGCACACCCCTTCTGCTGACTGACCCGACTACCTGCTCACACAGCTCCCG 1226  
QY 381 ProThrAlaPro--GlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIle 399  
Db 1227 CTTACAGCTCCAGTGTGGCCGCTCCTCACTGCTGCTCAACGCTGCTCAACGCTCCCTGCGCATC 1286  
QY 400 SerSerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysCys 419  
Db 1287 AGCAGTAGCTAGTGGCCATGCAACAGAGGCCCTTACGACGGGGCCCTGTGGCTCTGCTGC 1346  
QY 420 SerSerCysLeuAsnIleGlySerLysGlyLysSerSerTrpCysSerGluProThrLeu 439  
Db 1347 TCCAGCTGCTTAATATTGGGAACAGGGAAGTCTTCTACTGCTCTGAGCCTACCTC 1406  
QY 440 LeuLysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAla 459  
Db 1407 TTCAGAAAGGCCAATAAATAACAGAACTTCAAGTCAGAGGTGATGTACAGTGAAGGCT 1466  
QY 460 SerGlySerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476  
Db 1467 GTGGGACAGACCCACTGCGGGGTGGGACACTGGGGTGTCTCTCGTTATG 1517  
RESULT 7  
AF428262 2427 bp mRNA linear ROD 07-NOV-2001  
LOCUS Mus musculus calcium/calmodulin-dependent protein kinase I gamma  
DEFINITION (Camk1g) mRNA, complete cds.  
ACCESSION AF428262  
VERSION AF428262.1 GI:16755793  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 2427)  
AUTHORS Bjork, B.C., Watanabe, Y., Murray, J.C. and Schutte, B.C.  
TITLE Characterization of the human ortholog of rat Cam Kinase I gamma  
JOURNAL (Camk1g) at lq32-q41  
REFERENCE 2 (bases 1 to 2427)  
AUTHORS Bjork, B.C., Watanabe, Y., Murray, J.C. and Schutte, B.C.  
TITLE Direct Submission  
JOURNAL Submitted (05-OCT-2001) Genetics, University of Iowa, 140 EMRB,  
Iowa City, IA 52242, USA  
FEATURES  
source 1. 2427  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="1"  
/map="1H, 104.00 cM"

```
/clone="IMAGE:4502479"
/tissue_type="retina"
/notes="corresponds to EST B6293660"
1..2427
/gene="Camk1g"
86..1519
/gene="Camk1g"
/codon_start=1
/product="calcium/calmodulin-dependent protein kinase I
gamma"
/protein_id="AAL28101.1"
/db_xref="GI:16755794"
/translation="MGRKEEDCSKKKQTTNIRKTFIMFVLGSGAFSEVLYKQRV
TKGLFALKIKKSPAPRDSLENEIAVLAKRIKHENIVILEDIESTHYLYVMQLVSG
GEUFLRILKDYVTEKDALSLVIQQVLSAVKYLVHENGIVHRDLKPNLKYLPENSKI
MITDFGLSKMEQNGVMTACGTPGYVAPEVLAQPKYSKAVDCWSIGVATYILLCGYPP
FYETESKLFEDIKEGYEFESFPWDDISESAFDFICHLLEKDPNERTYCEKALRHWP
IDGNTALHRDIYPSVSLQIQKNFAKQWROAFNAAAVVHMKLHMNHSVSRQVEE
NRPPVSPAPVEVRPDSHDSITFAPILDPSTPLPALTRLPCSHSRPSPAPSGRSLNC
LVNGSIRISSLYPMQGPPLATGPGCCSSCLNIGNKSKSYCSYCEPTLFRKANKQNF
KSEVWVVRAGGSTHCRGGGTGVCVLM"
BASE COUNT 584 a 694 c 601 g 547 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 6.27e-157 Length: 2427
Score: 2327.50 Matches: 439
Percent Similarity: 94.97% Conservative: 14
Best Local Similarity: 92.03% Mismatches: 23
Query Match: 92.62% Indels: 1
DB: 10 Gaps: 1

US-09-960-643-2 (1-476) x AF428262 (1-2427)
```

```
Qy 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProProPheTyr 220
Db 686 GACTGCTGGTCCATTGGTGTTCATCACATACATACATGCTGTGTGGCTATCCCCCTTTCTAT 745

Qy 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
Db 746 GAAGAAACAGAAATCAAGAGCTTTTGAAGAATCAAGAGGGGTACTACGAGTTTGAGTCT 805

Qy 241 ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
Db 806 CCGTCTCGGATGACATTTCTGAGTCAGCCCAAGGATTTATTGGCCATCTGCTGGAGAAG 865

Qy 261 AspProAsnGluAurTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280
Db 866 GACCCAAACGAAACGGTACACCTGCGAGAAAGCCCTCAGACACCCCTGGATGTGTAAGAAC 925

Qy 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300
Db 926 ACAGCCCTGACCGGGACATCTATCCATCTGTACGCTCCAGATTCAGAGAACTTTGCC 985

Qy 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu 320
Db 986 AAGAGCAAGTGGAGGCAAGCCTTCAACGAGCTGCGGTGTGTCATCATACATGAGAAACTG 1045

Qy 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProGluThr 340
Db 1046 CACATGAACCTGCACAGCCCGCAGTGTCCGCCAAGAGGTGGAGACAGCCACCTGTGTCC 1105

Qy 341 GlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaProVal 360
Db 1106 CTGCTCTCTGAAGTCTCCAGACAGACTCCCATGACAGCTCCATCAGAGGGCCCCATC 1165

Qy 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380
Db 1166 GTGGACCCCAAGCACACACCCCTTCCTGTCACCTACCCGACTACCTGCTCACACAGCTCCGG 1225

Qy 381 ProThrAlaPro---GlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIle 399
Db 1226 CCTCAGCTCCCGAGTGGTGGCCGCTCACTCAACTGCCCTGGTCAACGGCTCCCTCGCGCATC 1285

Qy 400 SerSerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysCys 419
Db 1286 AGCAGTAGCTAGTGCCTATCAACAGAGGGCCCCCTAGCCAGGGGCGCTGTGGCTGTGC 1345

Qy 420 SerSerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCysSerGluProThrLeu 439
Db 1346 TCCAGCTGCTCTAAATATTGGGAACAAGGAAAGTCTTCTACTGCTGTGAGCCTACCCCTC 1405

Qy 440 LeuLysLysAlaAsnLysLysGlnAsnPhelLysSerGluValMetValProValLysAla 459
Db 1406 TTCAGAAAGGCCAATAAATAACACAACTTCAAGTCAGAGGTGATGGTACCAGTGAAGGCT 1465

Qy 460 SerGlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476
Db 1466 GTGTGCAGCAGCCCACTGCGGGGTGGGAGAGTGGGTGTCTCTCGTTATG 1516

RESULT 8
D86557
LOCUS D86557 Rattus norvegicus mRNA for Protein Kinase, partial cds.
DEFINITION D86557
ACCESSION D86557
VERSION D86557.1 GI:2077933
KEYWORDS Protein Kinase.
SOURCE Rattus norvegicus embryo (E18) brain cDNA to mRNA, clone_lib.S.
Nakanishi clone:N5.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1013)
AUTHORS Yokokura,H., Terada,O., Naito,Y. and Hidaka,H.
TITLE Isolation and comparison of rat cDNAs encoding
```

```
101 MetGlyArgLysGluGluAspAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
Db 86 ATGGGGCGTAAGGAGGAGGAGGACTGCAGTTCCTGGAAGAAACAGACACCAACATCAGG 145

Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 146 AAAACCTTCATCTTCATGGAAGTGTGGGATCAGGAGCTTCTCAGAGGTGTTCCTGGTG 205

Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
Db 206 AAGCAAGAGCTGACTGGAAACCTTTGCTCTGAATGATCAAGAAGTCACCAGCCTTC 265

Qy 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
Db 266 CGGACAGCAGCCCTAGAGAATGAGATCGCTGTGTTGAAAGGATCAAGCATGAGAACTT 325

Qy 81 ValThrLeuGluAspIleTyrGluSerThrHisTyrTyrLeuValMetGlnLeuVal 100
Db 326 GTGACCTTGAGGACATCTATGAGAGCACCCACCTACTACTCTGCTATGACGCTTGT 385

Qy 101 SerGlyGlyLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
Db 386 TCTGGAGGTGAGCTTTTGTACCGGATCCTAGAGCGTGTGCTACACAGAAAAAGGATGCC 445

Qy 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
Db 446 AGCCTGGTCATCCAGAGGCTTGTGTGCGGTGAATACTTCCATGAGAAATGGCATCGTC 505

Qy 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
Db 506 CACAGAGATCTAAAGCCTGAAACCTGCTGCTACCTCACCCCTGAGGAGAACTCCAAGATC 565

Qy 161 MetIleThrAspPheGlyLeuSerLysMetGluGluAsnGlyIleMetSerThrAlaCys 180
Db 566 ATGATCACTGACTTTGTGCTATCCAGATGGAGCAATGGAGTCACTGTCACAGCTTGT 625

Qy 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Db 626 GGGACCCAGGCTACGTGGCTCCAGAAAGTCTGCGCCAGAGAGCCCTACAGTAGGCTGTG 685
```



Qy	21	LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal	40
Db	61	AGATCTTCGAGTTCAAGAGACCCCTCGGAACGGGGCGCTTTCGGAAGTGGTTAGCT	120
Qy	41	LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle--LysLysSerProAla	59
Db	121	GAAGAGAGGCNACTGGCAAGCTCTTGTGCTGAAGTGTATCCCTAGAAGGGCGCTGAAG	180
Qy	60	PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn	79
Db	181	GGCAAGGAAAGCAGCATAGAGAATGAGATGCGCTCTCGAAGAGATTAAAGCATGAAAAAT	240
Qy	80	IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu	99
Db	241	ATTGTGCCCTGGAGACATTTATGAAGCCCAATCACTGTACTTGGTCATGCGACGTG	300
Qy	100	ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp	119
Db	301	GTCTCGGTGGAGAGCTGTGTGACCGGATAGTGGAGAAAGGGGTTTATACAGAGAAAGAT	360
Qy	120	AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle	139
Db	361	GCAGACACTGTATCGCGCAAGCTCTTGGACGGCGGTACTATCTCCAGAAATGGGCATC	420
Qy	140	ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys	159
Db	421	GTCCACAGAGACCTCAAGCCCGAATCTCTTGACTACAGTCAGATGAGGAGTCCAAA	480
Qy	160	IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly--IleMetSerThr	178
Db	481	ATAATGATCAGTGATTTGGATTGTCAAAATGGAGGCAAGGAGATGTGATGCCACT	540
Qy	179	AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys	198
Db	541	GCCTGTGNACTCCAGCCTATGTCGCTCTCGAGTCTCGCCAGAAACCTTACAGCANA	600
Qy	199	AlaValAspCysTyrSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro	218
Db	601	GCGGTGTACTGCTGCTCATCGGAGTGATGCTCATCTTGTCTGCTGCGGTACCCCTCT	660
Qy	219	PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe	238
Db	661	TTTATATGATGAATGACTCCAAGCTTTTGACGAGATCCTCAAGCGCGAATATGAGTTT	720
Qy	239	GluSerProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu	258
Db	721	GACTCTCCCTACTGGGATGACATCTCCGACTGTGAAAAGACTTCATTTCGGAACCTGATG	780
Qy	259	GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTyrIleAsp	278
Db	781	GAGAGGACCCGAATAAAGATACACGTGTGACGAGGACGCTCGGCCACCCATGGATCGCT	840
Qy	279	GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn	298
Db	841	GGTGACACAGCCCTCAACAAAAACATCCACGAGTCCGTCCGCGCCGACATCCGGAANAAC	900
Qy	299	PheAlaLysSerLysTyrArgGlnAlaPheAsnAlaAlaValValHisHisMetArg	318
Db	901	TTTGCCACAGACCAATGGAGACAGCATTTAATGCCACGGCGCTGCTGAGACATATGAGA	960
Qy	319	LysLeuHisMet	322
Db	961	AAACTACACCTC	972

RESULT 10				
AXI67587	AXI67587	1074 bp	DNA	linear
LOCUS	Sequence 3 from Patent WO0142435.			
DEFINITION	AXI67587			
ACCESSION	AXI67587.1	GI:14596987		
VERSION	.			
KEYWORDS	human.			
SOURCE	Homo sapiens			
ORGANISM				

```
Qy 239 GluSerProPheTrpAspPheIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 721 GACTCCCTACTGGATGACATCTCCGACTTCGCAAAAGACTTATTCGGAACTGATG 780

Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpLeuAsp 278
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 781 GAGAAGGACCCGAATAAAGATACAGCTGTGAGCAGGAGCTCGGCACCATGGATCGCT 840

Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 841 GGTGACACAGCCCTCAACAAACATCCACGAGTCGGTCAGCCCGCAGATCCGGAAGAAC 900

Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 901 TTTGCCAAGACAAATGGAGACAGCATTTAATGCCACGCCGCTCGTGAGACATATGAGA 960

Qy 319 LysLeuHisMet 322
|||||:::|||||:::
Db 961 AACTACACCTC 972

RESULT 11
LOCUS AX167585 1158 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 1 from Patent WO0142435.
ACCESSION AX167585
VERSION AX167585.1 GI:14596986
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1158)
AUTHORS Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B.,
Abuin,A. and Sands,A.T.
TITLE Novel human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0142435-A 1 14-JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1. .1158
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 323 a 283 c 304 g 248 t
ORIGIN

Alignment Scores:
Pred. No.: 3 94e-80 Length: 1158
Score: 1246.00 Matches: 237
Percent Similarity: 85.49% Conservative: 40
Best Local Similarity: 73.15% Mismatches: 45
Query Match: 49.58% Indels: 2
DB: Gaps: 2

US-09-960-643-2 (1-476) x AX167585 (1-1158)

Qy 1 MetGlyArgLysGluGluAspAspCysSerTrpLysGlnThrThrAsnIleArg 20
||| |||::: ||| |||::: ||| |||::: ||| |||::: ||| |||::: ||| |||:::
Db 1 ATGCGCGGAGAACGCGAGAGCAGCTCTCTCTGGAAGAACGAGCTGAAGACATCAAG 60

Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AAGATCTTCGAGTTCGAAGACCCCTCGGAACCGGGGCCCTTTCCGAAGTGGTTAGCT 120

Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59
:::|||||::: ||| |||::: ||| |||::: ||| |||::: ||| |||::: ||| |||:::
Db 121 GAAGAGAAGCAACTGGCAAGCTCTTTGCTGTGAAGTGTATCCCTAAGAAGCGCTGAAG 180

Qy 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
:::|||||::: ||| |||::: ||| |||::: ||| |||::: ||| |||::: ||| |||:::
Db 181 GGCAAGGAAGAACGACATAGAGAATGAGATAGCGCTCTGAGAAGGATTGAAGCATGAAAT 240

Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 ATTGTTCCCTGGAGACATTTATGAAGCCCAATCACCTGTACTTGTGTCATGACGCTG 300
```

```
Qy 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 301 GTGTCCCGTGGAGAGCTGTTTGGACCGATAGTGGAGAAGGGGTTTATACAGAAGGAT 360

Qy 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 361 GCCAGCACTCTGATCGCCCAAGTCTTGGACGCGGTACTATCTCCACAGAAGTGGGCATC 420

Qy 140 ValHisArgAspLeuLysProGluAsnLeuTyrLeuThrProGluGluAsnSerLys 159
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 421 GTCCACAGAGACCTCAAGCCCGAAATCTCTGTACTACAGTCAAGATAGAGAGTCCAAA 480

Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGlnAsnGly---IleMetSerThr 178
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 481 ATAATGATCAGTACTTTGGATTGTCAAAATGGAGGCGCAAGAGATGTGATGTCCTCACT 540

Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 541 GCCTGTGGAACCTCCAGGCTATGTCCTCGTGAAGTCTCGCCAGAGAACTTACAGCAA 600

Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 601 GCCGTTGACTGCTGGTCCATCGGAGTATGCTTACATCTTGCTCTCGGCTACCCCTCT 660

Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 661 TTTTATGATGAATGACTCCAGCTCTTTGAGCAGATCTCAAGGCGGGAATATGATTT 720

Qy 239 GluSerProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 721 GACTCTCCCTACTGGATGACATCTCGGACTCTCGAAAGACTTCAATCGGAACCTGATG 780

Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 781 GAGAAGGACCCGAATAAAGATACACGCTGTGAGCAGCAGCTCGGACCCATGATCGCT 840

Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 841 GGTGACACAGCCCTCAACAAACATCCACGAGTCCGTCAGCGGCCAGATCCGGGAAAC 900

Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValAlaValHisHisMetArg 318
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 901 TTTGCCAAGACAAATGGAGACAGCATTTAATGCCACGCCGCTCGTGAGACATATGAGA 960

Qy 319 LysLeuHisMet 322
|||||:::|||||:::
Db 961 AACTACACCTC 972

RESULT 12
LOCUS AF286366 1579 bp mRNA linear PRI 17-AUG-2000
DEFINITION Homo sapiens CamKI-like protein kinase mRNA, complete cds.
ACCESSION AF286366
VERSION AF286366.1 GI:9837340
KEYWORDS .
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1579)
AUTHORS Verploegen,S., Koenderman,L. and Coffey,P.J.
TITLE Identification and characterization of CKIIk: a novel granulocyte
JOURNAL Blood (2000) In press
REFERENCE 2 (bases 1 to 1579)
AUTHORS Verploegen,S. and Coffey,P.J.
TITLE Direct Submission
JOURNAL Submitted (11-JUL-2000) Dept. Pulmonary Diseases, University
Medical Center Utrecht, Heidelberglaan 100, Utrecht 3584 CX, The
Netherlands
FEATURES
Location/Qualifiers
source 1. .1579
```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="eosinophil"
88..1161
/notes="CKL1k"
/codon_start=1
/product="CamK1-like protein kinase"
/protein_id="AAG00534.1"
/db_xref="GI:9837341"
/translation="MARENGSSSWKKOAEIDIKIPFEKTLGTGAPSEVLAEEK
TKGLFAVKCIPKALKGKSESTIENIAVRLKIHENI VALEDIYESPHLYLVQLVS
GGELDRIVEKGFYTKDASTLIRQVDAVYLHRWGIYHRLDLPENLIYISQDESK
IMISDFGLSKMGKGDVMTAGCTPGYVAPYLAQKPYSKAVDCWISGVATYLLCY
PFYDENSCLKFEQLKAEYDFSPYDDI SDSAKDFIRNLMEKDPNKRITCEQAAH
PWAGDTALNKIHESVSAQIRKNFPAKSKRQAFNATAVVRHMRKLLHGLSSLSNNA
VSSSLASLQKDCAYVAKPESL"
BASE COUNT 422 a 414 c 415 g 328 t
ORIGIN
Alignment Scores:
Pred. No.: 5.88e-80 Length: 1579
Score: 1246.00 Matches: 237
Percent Similarity: 85.49% Conservative: 40
Best Local Similarity: 73.15% Mismatches: 45
Query Match: 49.58% Indels: 2
DB: 9 Gaps: 2
US-09-960-643-2 (1-476) x AF286366 (1-1579)
Qy 1 MetGlyArgLysGluGluAspAspCysSerSerTrpLysGlnThrThrAsnIleArg 20
Db 88 ATGCCCGGGAGACGCGAGACGAGCTCTCTCTGGAAAAGCAAGCTGAGACATCAAG 147
Qy 21 LysThrPheIlePheMetGluValValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 148 AAGATCTTCAGATTCAAAGAGACCTCGGAACCGGGCCCTTTCCGAGTGGTTTTAGCT 207
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59
Db 208 GAAGAGAAGCAACTGGCAAGCTCTTGTGTGAAGTGTATCCCAAGAAGCGCGTGAAG 267
Qy 60 PheArgAspSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
Db 268 GCGAAGAAAGACGACATGAGATGAGATAGCGCTCTCGAGAAGATTGAAGATGAAAT 327
Qy 80 IleValThrLeuGluAspIleTyrgluSerThrThrHisTyrrTyrrLeuValMetGlnLeu 99
Db 328 ATTGTGCTGGAGACATTTATGAAGCCCAATCACCTGTACTGTGTCATGACGCTG 387
Qy 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrrThrGluLysAsp 119
Db 388 GTGTCCGGTGGAGAGCTGTTTGACCGGATAGTGGAGAGGGGTTTATACAGAGAAGAT 447
Qy 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValIlyTyrrLeuHisGluAsnGlyIle 139
Db 448 GCCAGACTCGATCCGCCAAGTCCTGGACGCCGTGTACTATCTCCACAGATGGGCATC 507
Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrrLeuThrProGluGluAsnSerLys 159
Db 508 GTCCACAGAGACCTCAAGCCCGAAATCTCTTGTACTACAGTCAAGTAGGAGTCCAAA 567
Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178
Db 568 ATAATGATCAGTGACTTGGATTGTCAAAAATGGAGGGCAAGAGGATGTGATGTCCACT 627
Qy 179 AlaCysGlyThrProGlyTyrrValAlaProGluValIleAlaGlnLysProTyrrSerLys 198
Db 628 GCCTGTGGAACTCCAGCGCTATGCTGCTCTCTGAAGTCTCTGCCCCAGAAACCTTACACAAA 687
Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrrIleLeuLeuCysGlyTyrrProPro 218
Db 688 GCCCTGTGACTGCTGGTCCATCGGAGTGATTGCCTACATCTTGTCTGCGGCTACCCCTCT 747
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="eosinophil"
88..1161
/notes="CKL1k"
/codon_start=1
/product="CamK1-like protein kinase"
/protein_id="AAG00534.1"
/db_xref="GI:9837341"
/translation="MARENGSSSWKKOAEIDIKIPFEKTLGTGAPSEVLAEEK
TKGLFAVKCIPKALKGKSESTIENIAVRLKIHENI VALEDIYESPHLYLVQLVS
GGELDRIVEKGFYTKDASTLIRQVDAVYLHRWGIYHRLDLPENLIYISQDESK
IMISDFGLSKMGKGDVMTAGCTPGYVAPYLAQKPYSKAVDCWISGVATYLLCY
PFYDENSCLKFEQLKAEYDFSPYDDI SDSAKDFIRNLMEKDPNKRITCEQAAH
PWAGDTALNKIHESVSAQIRKNFPAKSKRQAFNATAVVRHMRKLLHGLSSLSNNA
VSSSLASLQKDCAYVAKPESL"
BASE COUNT 422 a 414 c 415 g 328 t
ORIGIN
Alignment Scores:
Pred. No.: 5.88e-80 Length: 1579
Score: 1246.00 Matches: 237
Percent Similarity: 85.49% Conservative: 40
Best Local Similarity: 73.15% Mismatches: 45
Query Match: 49.58% Indels: 2
DB: 9 Gaps: 2
US-09-960-643-2 (1-476) x AF286366 (1-1579)
Qy 1 MetGlyArgLysGluGluAspAspCysSerSerTrpLysGlnThrThrAsnIleArg 20
Db 88 ATGCCCGGGAGACGCGAGACGAGCTCTCTCTGGAAAAGCAAGCTGAGACATCAAG 147
Qy 21 LysThrPheIlePheMetGluValValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 148 AAGATCTTCAGATTCAAAGAGACCTCGGAACCGGGCCCTTTCCGAGTGGTTTTAGCT 207
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59
Db 208 GAAGAGAAGCAACTGGCAAGCTCTTGTGTGAAGTGTATCCCAAGAAGCGCGTGAAG 267
Qy 60 PheArgAspSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
Db 268 GCGAAGAAAGACGACATGAGATGAGATAGCGCTCTCGAGAAGATTGAAGATGAAAT 327
Qy 80 IleValThrLeuGluAspIleTyrgluSerThrThrHisTyrrTyrrLeuValMetGlnLeu 99
Db 328 ATTGTGCTGGAGACATTTATGAAGCCCAATCACCTGTACTGTGTCATGACGCTG 387
Qy 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrrThrGluLysAsp 119
Db 388 GTGTCCGGTGGAGAGCTGTTTGACCGGATAGTGGAGAGGGGTTTATACAGAGAAGAT 447
Qy 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValIlyTyrrLeuHisGluAsnGlyIle 139
Db 448 GCCAGACTCGATCCGCCAAGTCCTGGACGCCGTGTACTATCTCCACAGATGGGCATC 507
Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrrLeuThrProGluGluAsnSerLys 159
Db 508 GTCCACAGAGACCTCAAGCCCGAAATCTCTTGTACTACAGTCAAGTAGGAGTCCAAA 567
Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178
Db 568 ATAATGATCAGTGACTTGGATTGTCAAAAATGGAGGGCAAGAGGATGTGATGTCCACT 627
Qy 179 AlaCysGlyThrProGlyTyrrValAlaProGluValIleAlaGlnLysProTyrrSerLys 198
Db 628 GCCTGTGGAACTCCAGCGCTATGCTGCTCTCTGAAGTCTCTGCCCCAGAAACCTTACACAAA 687
Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrrIleLeuLeuCysGlyTyrrProPro 218
Db 688 GCCCTGTGACTGCTGGTCCATCGGAGTGATTGCCTACATCTTGTCTGCGGCTACCCCTCT 747
Qy 219 PheTyrrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrrTyrrGluPhe 238
Db 748 TTTTATGATGAAATGACTCCAGCTCTTTGAGCAGATCTCTCAAGGCGGAATATGAGTTT 807
Qy 239 GluSerProPheTrpAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
Db 808 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAGACTTCAATTCGGAAACCTGATG 867
Qy 259 GluLysAspProAsnGluArgTyrrCysGluLysAlaLeuSerHisProTrpIleAsp 278
Db 868 GAGAAGACCCGAAATAAAGATACAGCTGTGAGCAGGACGCTCGGACCCATGGATCGCT 927
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrrProSerValSerLeuGlnIleGlnLysAsn 298
Db 928 GGTGACACAGCCCTCAACAAAAACATCCACGAGTCCGTGAGCGCCAGATCCGGAAGAAC 987
Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318
Db 988 TTGCCCCAAGACCAANTGGAGACAGCATTTAATGCCACGCCGCTGCTGAGACATATGAGA 1047
Qy 319 LysLeuHisMet 322
Db 1048 AAATACACCTC 1059
RESULT 13
AX167589 LOCUS AX167589 1671 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 5 from Patent WO0142435.
ACCESSION AX167589
VERSION AX167589.1 GI:14596988
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1671)
AUTHORS Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B.,
Abuin,A. and Sands,A.T.
TITLE Novel human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0142435-A 5 14-JUN-2001.
LEXICON Lexicon Genetics Incorporated (US)
FEATURES
source
location/Qualifiers
source
1..1671
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 453 a 437 c 448 g 333 t
ORIGIN
Alignment Scores:
Pred. No.: 6.33e-80 Length: 1671
Score: 1246.00 Matches: 274
Percent Similarity: 64.35% Conservative: 60
Best Local Similarity: 52.79% Mismatches: 97
Query Match: 49.58% Indels: 91
DB: 6 Gaps: 11
US-09-960-643-2 (1-476) x AX167589 (1-1671)
Qy 1 MetGlyArgLysGluGluAspAspCysSerSerTrpLysGlnThrThrAsnIleArg 20
Db 167 ATGCCCGGGAGAACGCGAGACGAGCTCTCTCTGGAAAAGCAAGCTGAAAGACATCAAG 226
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 227 AAGATCTTCGAGTTCAAAGACCCCTCGGAACCGGGCCCTTTCCGAAGTGGTTTAGCT 286
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59
Db 287 GAAGAGAAGCAACTGGCAAGCTCTTTGCTGTGAGTGTATCCCTAAGAAGCGGCTGAAG 346
Qy 60 PheArgAspSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
Db 347 GCGAAGAAAGACGACATGAGATGAGATAGCGCTCTCGAGAAGATTAAGCATGAAAT 406
```

```
Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99
Db 407 ATTTGTCCTCGGAGACATTTATGAAGCCCAATCAGCTGTACTGTGCTATGCTGAGCTG 466
Qy 100 ValSerGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119
Db 467 GTGTCCGGTGGAGAGCTGTGTGACGGATAGTGGAGAAGGGTTTATACAGAGAGGAT 526
Qy 120 AlaSerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139
Db 527 GCCAGCACTGTATCCGCCAAGTCTTGGACGCGTGTACTATCTCCACAGAATGGGCATC 586
Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159
Db 587 GTCCACAGAGACCTCAAGCCGAAAATCTCTGTACAGTCAAGATGAGGAGTCCAAA 646
Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178
Db 647 ATAATGATCAGTGCATTTGGATTGTCAAAAATGGAGGGCAAGGATGTGATGTCCACT 706
Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198
Db 707 GCCTGTGGAACTCCAGCTATGCTGCTCTGAAAGTCTGCCCGAGAAACCTTACAGCAA 766
Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
Db 767 GCCGTTGACTGCTGGTCCATCGAGTATGTCCTACATCTGCTCTGCGGCTACCTCTCT 826
Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238
Db 827 TTTTATGATAAATGATGATCCCAAGCTCTTTGACAGATCTCAAGCGCGAAATATGAGTTT 886
Qy 239 GluSerProPheTrpAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
Db 887 GACTCTCCCTACTGGGATGACATCTCGACTCTGCAAAAGACTTCATTCGGAACCTGAG 946
Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278
Db 947 GAGAAGGACCCGATAAAGATACAGCTGTGAGCAGGAGCTCGGCACCCATCGATCGCT 1006
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
Db 1007 GGTACACAGCCCTCAACAAAACATCCAGAGTCCGTCAGCGCCCGATCGCGAAAC 1066
Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318
Db 1067 TTTGCCAAGCAAAATGGAGACAAGCATTTAATGCCAGCGCTCGTGAGACATATGAGA 1126
Qy 319 LysLeuHisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArg----- 336
Db 1127 AAATACACCTCGCGAGCAGCCCTGGACAGTTCAAATGCAAGTGTTCGAGCAGCCCTCAGT 1186
Qy 337 ---ProProGluThrGlnAlaSerGluThrSerArgProSerSerProGluIleThrIle 355
Db 1187 TTGCCAGCCCAAAAGACTGTGCTATGTAGTGAACAAACAGAAATC-CCTCAGCTGACACTG 1245
Qy 356 ThrGluAlaProValLeuAspHisSerValAlaLeuProAla----- 369
Db 1246 AA-GACGAGCTGCTGGGTGGAGAGGAGGAGCGCGGCATCTGCCGAGCACCTCCTGTTGCC 1304
Qy 370 -----LeuThrGlnLeuProCys----- 375
Db 1305 AGGCGCTTTCTATACTTAATCCCATGTCATCGCACCTAGGACTTTTTTTTACATGTAAT 1364
Qy 376 -----GlnHis---GlyArg----- 379
Db 1365 CACTGGGCTGGGTGCAGTGGCTCACGCTGTATATCCCAACACTTCTGGGAGCTGAGGCA 1424
Qy 380 -----ArgProThrAlaPro----- 384
Db 1425 GAGGACTGTTGAGTTCAGGAGTTTAAAGACCAAGCTTGACCAACATGTTGAACCCCATC 1484
```

```
Qy 385 -----GlyGlyArgSerLeuAsnCysLeuValAsn 394
Db 1485 TCTCTAAATATAATAATTAGCGGTGGTGGCGAGCACCTGTGAATGTCAGCTACTT 1544
Qy 395 GlySerLeuHisIleSerSerSerLeuValProMetHisGlnGlySerLeuAlaAlaGly 414
Db 1545 GGGAGGCTGAGCCA-GGAGAATCGCTTGAGCCC-----GGA 1579
Qy 415 ProCysGlyCysCysSerSerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCys 434
Db 1580 AGCGAGGTTGC-----AATGAGCTGAGATCACACCACTGC 1615
Qy 435 Ser-GluPro-----ThrLeuLeuLysLysAlaAsnLysLysGln 447
Db 1616 ACTTCAGCTGGTGCACAGATTGAGACTCTCTCAAAAAAATAAAAAA 1670

RESULT 14
BC014825 1448 bp mRNA linear ROD 07-AUG-2002
LOCUS BC014825
DEFINITION Mus musculus, Similar to calcium/calmodulin-dependent protein
kinase I, clone MGC:18933 IMAGE:3969343, mRNA, complete cds.
ACCESSION BC014825
VERSION BC014825.1 GI:15928725
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1448)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapsb-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 23 Row: j Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
FEATURES
source Location/Qualifiers
1..1448
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:18933 IMAGE:3969343"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha
model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI_CGAP_Mam1"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
120..1244
/codon_start=1
/product="Similar to calcium/calmodulin-dependent protein
kinase I"
/protein_id="AAH14825.1"
/db_xref="GI:15928726"

CDS
```

/translation="MPGAVEGPRWQAEIRDYDFRDVLGTGAFSEVILAEKRTQK  
LVAIKCIAKALEGKSGMENEIAVLHKIKHPNIVALDDIYESGGHLYLIMQLVSGGE  
LFDRIVEKGYTERDASRLIFQVLDVAVKYLHDLGIVHRLKPNLKYSLDEDSKIMI  
SDFGLSKMEDPGSVLSTACGTGYYVAPEVLAOKPYSKAVDCWSIGVIAIYILLCGYPPF  
YDNDAKLFEQILKAEYEPDPSYDDIISDSAKDFIRHLMKDPKRFCEQALQHPWI  
AGTALDKNIHQSVSQIKKNFAKSKWQAFNATAVVRHMRKLQLTSGEQGGQTGSH  
GELLITPAGGPAAGCCCRDCCVEPSELPAPPSSRAMD"

BASE COUNT 364 a 375 c 408 g 301 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1,19e-73 Length: 1448  
Score: 1157.00 Matches: 237  
Percent Similarity: 70.94% Conservative: 51  
Best Local Similarity: 58.37% Mismatches: 64  
Query Match: 46.04% Indels: 54  
DB: 10 Gaps: 8

US-09-960-643-2 (1-476) x BC014825 (1-1448)

Qy 12 TrpLysLysGlnThrThrAsnIleArgLysThrPheIlePheMetGluValLeuGlySer 31  
Db 147 TGG---AAGCAGCGGGAAGACATTAGGATATTTATGACTTTCAGAGATGTTCTGGGCACG 203  
Qy 32 GlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGlyLysLeuPheAlaLeu 51  
Db 204 GTGCTCTTCAGAAAGTGATCCTGGCAGGACACAGAGACTCAGAACTGGTGGCCATC 263  
Qy 52 LysCysIle---LysLysSerProAlaPheArgAspSerSerLeuGluAsnGluIleAla 70  
Db 264 AAATGCATTGCCAAGAAAGCCCTGGAGGGCAAGAGGAGCAGATGCGAGAACGAGATTGCC 323  
Qy 71 ValLeuLysLysIleLysLysHisGluAsnIleValThrLeuGluAspIleThrGluSerThr 90  
Db 324 GTCTTACACAGATCAAGACCCCAACATGTTAGCCCTGGATGACATCTATGAGATGGG 383  
Qy 91 ThrHisThrLeuValMetGlnLeuValSerGlyGlyGluLeuPheAspArgIleLeu 110  
Db 384 GCCACCTCTACCTCATCATGAGCTGGTTTTCAGGTGGAGAGCTGTTTACCCGAATTGG 443  
Qy 111 GluArgGlyValThrThrGluLysAspAlaSerLeuValIleGlnValLeuSerAla 130  
Db 444 GAGAAAGGATTCTACACGGAAACGGGATGCCAGCGCGCTCATCTCCAGGTGGTGGATGCT 503  
Qy 131 ValLysThrLeuHisGluAsnGlyIleValHisArgAspLeuLysProGluAsnLeu 150  
Db 504 GTCAGTACCTGACGACCTGGGCTGCTCCACAGCCTGGGACTCCAGGATATGCGCCCTGAG 563  
Qy 151 TyrLeuThrProGluGluAsnSerLysIleMetIleThrAspPheGlyLeuSerLysMet 170  
Db 564 TACTACAGTCTGGATGAAGACTCCAAATCATGATCTCTGACITTTGGCCTCTCCAGATG 623  
Qy 171 GluGlnAsnGly---IleMetSerThrAlaCysGlyThrProGlyThrValAlaProGlu 189  
Db 624 GAGACCCAGCAGCAGTGTGCTCTCCACAGCCTGGGACTCCAGGATATGCGCCCTGAG 683  
Qy 190 ValLeuAlaGlnLysProThrSerLysAlaValAspCysThrPheIleGlyValIleThr 209  
Db 684 GTCTGCGCCAGAGCCCTACAGCAGGCTGTGATGTTGCTGTCATAGAGTCATGCTGCC 743  
Qy 210 TyrIleLeuLeuCysGlyTyrProPheThrGluGluThrGluSerLysLeuPheGlu 229  
Db 744 TATATGCTGCTCTGTGGCTACCCACCCTTTTATGATGATAAATGATGCCAAACTCTTTGAA 803  
Qy 230 LysIleLysGluGlyTyrThrGluPheGluSerProPheThrPaspAspIleSerGluSer 249  
Db 804 CAGATTGTAAGGACAGATGATGATGACTCTCCCTATTGGGACACATCTCTGACTCT 863  
Qy 250 AlaLysAspPheIleCysHisLeuLeuGluLysAspProAsnGluArgThrThrCysGlu 269  
Db 864 GCCAAAGATTTCATACACATTTGATGGAGAAAGACCCGGAGAGAGGTTTACCTGTGAG 923  
Qy 270 LysAlaLeuSerHisProThrIleAspGlyAsnThrAlaLeuHisArgAspIleThrPro 289

Db 924 CAGGCTTGCAGCACCCATGATGTCAGAGATACACTCTGGATAAGAAATATCCACCAG 983  
Qy 290 SerValSerLeuGlnIleGlnLysAsnPheAlaLysSerLysThrPheGlnAlaPheAsn 309  
Db 984 TCAGTGACGAGCAGATCAAGAGAACCTTTCAGGAGCAGGTCGAGGAGCTTTCAT 1043  
Qy 310 AlaAlaValValHisHisMetArgLysLeuHisMetAsnLeuHisSerProGly--- 328  
Db 1044 GCCACTGTGCTGTGTCGACACATCAGGAAGCTGCGGACCACTCAGGAGGGGCAG 1103  
Qy 328 ----- 328  
Db 1104 GGCAGACTGCGACGCCAGCGGAGCTGCTGACACCAACAGCTGCTGGGCCAGAGTGGC 1163  
Qy 329 -----ValArgProGluValGlu-----AsnArgProPro 338  
Db 1164 TGCTGCTGCCGAGACTGCTGTGTGGAACCAAGCTCGGAACTGCCCTCGCACCCCA 1223  
Qy 339 GluThrGlnAlaSerGlu---ThrSerArgProSer----- 349  
Db 1224 AGCTTAGGCGCATGATGTAACATCAAGGAAGTCAGGTTTCTTTCATGAGAGGGTTG 1283  
Qy 350 -----SerProGluIle-----ThrIleThrGluAla 358  
Db 1284 GAGCAGCCTGCTTCTCTCTCCTCCTCAGATCTGGGGGCTTCTGACCCCACTCTTCA 1343  
Qy 359 ProValLeuAspHisSer 364  
Db 1344 CCCCCTCCTCTCACTCT 1361  
RESULT 15  
RATCAMPKAA  
LOCUS Rattus norvegicus Cam-like protein kinase mRNA linear ROD 18-MAY-1995  
DEFINITION Rattus norvegicus Cam-like protein kinase mRNA, complete cds.  
ACCESSION L26288  
VERSION L26288.1 GI:439613  
KEYWORDS protein kinase  
SOURCE Rattus norvegicus (strain Sprague-Dawley) (library: day 21  
gestation lung cDNA) fetal lung cDNA to mRNA.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 1439)  
AUTHORS Cho,F.S., Phillips,K.S., Bogucki,B. and Weaver,T.E.  
TITLE Characterization of a rat cDNA clone encoding  
calcium/calmodulin-dependent protein kinase I  
JOURNAL Biochim. Biophys. Acta 1224 (1), 156-160 (1994)  
MEDLINE 95035115  
PUBMED 7948038  
FEATURES  
source Location/Qualifiers  
1. .1439  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/tissue\_type="lung"  
/dev\_stage="fetal"  
/disease\_lib="day 21 gestation lung cDNA"  
118. .1242  
/codon\_start=1  
/product="Cam-like protein kinase"  
/protein\_id="AAA66944.1"  
/db\_xref="GI:439614"  
/translation="MPGAVEGPRWQAEIRDYDFRDVLGTGAFSEVILAEKRTQK  
LVAIKCIAKALEGKSGMENEIAVLHKIKHPNIVALDDIYESGGHLYLIMQLVSGGE  
LFDRIVEKGYTERDASRLIFQVLDVAVKYLHDLGIVHRLKPNLKYSLDEDSKIMI  
SDFGLSKMEDPGSVLSTACGTGYYVAPEVLAOKPYSKAVDCWSIGVIAIYILLCGYPPF  
YDNDAKLFEQILKAEYEPDPSYDDIISDSAKDFIRHLMKDPKRFCEQALQHPWI  
AGTALDKNIHQSVSQIKKNFAKSKWQAFNATAVVRHMRKLQLTSGEQGGQTASH  
GELLITPAGGPAAGCCCRDCCVEPSELPAPPSSRAMD"  
BASE COUNT 357 a 375 c 405 g 302 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1,51e-73 Length: 1439  
Score: 1155.50 Matches: 249  
Percent Similarity: 64.66% Conservative: 51  
Best Local Similarity: 53.66% Mismatches: 85  
Query Match: 45.98% Indels: 80  
DB: 10 Gaps: 9

US-09-960-643-2 (1-476) x RATCAMPKAA (1-1439)

```
Qy 12 TriplysLysGlnThrThrAsnIleArgLysThrPheIlePheMetGluValLeuGlySer 31
Db 145 TGG---AAGCAGGACAGACATAGGATATTTATGACTTCAGGATGTTCTGGGCAGG 201
Qy 32 GlyAlaPheSerGluValPheLeuValLysGlnArgLysThrGlyLysLeuPheAlaLeu 51
Db 202 GGTGCCCTCTCAGAAGTGATCCTGGCAGAGGACAGAGGACTCAGAAACTGGTGGCCATC 261
Qy 52 LysCysIle---LysLysSerProAlaPheArgAspSerSerLeuGluAsnGluIleAla 70
Db 262 AATGCAATTCACAGAGGAGCCCTGGAGGCAAGAGAGGAGCATGAGATGAGATCGCC 321
Qy 71 ValLeuLysLysIleLysHisGluAsnIleValThrLeuGluAspIleTyrGluSerThr 90
Db 322 GTCTTACACAGATCAAGCACCACCATTTGATGCCCTGGATGACATCTATGAGAGTGGG 381
Qy 91 ThrHisTyrTyrLeuValMetGlnLeuValSerGlyGlyGluLeuPheAspArgIleLeu 110
Db 382 GGCACCTCTACCTCATCTGACGTGGTTCAGGTGGAGAGCTGTTGACCGAATGTG 441
Qy 111 GluArgGlyValTyrThrGluLysAspAlaSerLeuValIleGlnValLeuSerAla 130
Db 442 GAGAAGGATTCTACACAGAGGGATGCCAGCGCCTCATCTCCAGGTGCTGGATGCT 501
Qy 131 VallysTyrLeuHisGluAsnGlyIleValHisArgAspLeuLysProGluAsnLeuLeu 150
Db 502 GTCAAGTACCTGTCAGCAGCTGGGCGATTGTGCACCGGATCTCAAGCCAGAGACCTGCTA 561
Qy 151 TyrLeuThrProGluGluAsnSerLysIleMetIleThrAspPheGlyLeuSerLysMet 170
Db 562 TACTACAGTCTGGATGAAGACTCCAAATCATCATCTGACTTTGGCCTCTCCAAGATG 621
Qy 171 GluGlnAsnGly---IleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGlu 189
Db 622 GAGACCCAGGAGTGTGCTCTCCACAGCCTGTGGACTCCAGGATATGTGGCCCTGAG 681
Qy 190 ValLeuAlaGlnLysProTyrSerLysAlaValAspCysTrpSerIleGlyValIleThr 209
Db 682 GTCTTGCCCCAGAGCCCTACAGCAGGCGGTGGACTGTGGTCCATAGGATCATTTGCC 741
Qy 210 TyrIleLeuLeuCysGlyTyrProProPheTyrGluThrGluSerLysLeuPheGlu 229
Db 742 TATATACTGCTCTGTGGCTACCGCCGCTTATGATGAAATGATGCCAAACTTTTGA 801
Qy 230 LysIleLysGluGlyTyrTyrGluPheGluSerProPheTrpAspIleSerGluSer 249
Db 802 CAGATTTTGAAGGTGATGATGATTTGACTCTCCCTATTGGGACGACATCTCTGACTCT 861
Qy 250 AlaLysAspPheIleCysHisLeuLeuGluLysAspProAsnGluArgTyrThrCysGlu 269
Db 862 GCCAAGATTTTACATCCTATTTGATGGAGAACCCAGAGAGAGGTTACCTCTGAG 921
Qy 270 LysAlaLeuSerHisProTrpIleAspGlyAsnThrAlaLeuHisArgAspIleTyrPro 289
Db 922 CAGGCTTGGACGACCCCTGGATTCAGAGACACACAGCTCTGGATAAGATATCCACCAG 981
Qy 290 SerValSerLeuGlnIleGlnLysAsnPheAlaLysSerLysTrpArgGlnAlaPheAsn 309
Db 982 TCAGTGAGTGAGCAGATCAGAGAACTTTGCCAAGCAAGTGGAGCAAGCTTTCAAT 1041
Qy 310 AlaAlaAlaValValHisHisMetArgLysLeuHisMetAsnLeuHisSerProGlyVal 329
Db 1041 GCTACCGCTGTGGTTGGGCACATGAGGAAGCTGCAGCTGGGC----- 1083
Qy 330 ArgProGluValGluAsnArgProProGluThrGlnAlaSerGluThrSerArgProSer 349
Db 1084 -----ACGAGC----- 1089
Qy 350 SerProGluIleThrIleThrGluAlaProValLeuAspHisSerValAlaLeuProAla 369
Db 1089 ----- 1089
Qy 370 LeuThrGlnLeuProCysGlnHisGlyArgArgProThrAlaProGlyGlyArgSerLeu 389
Db 1090 -----CAGGAGGACAGGACAGACAGCTAGCCACGAGAGCTGCTG 1131
Qy 390 AsnCysLeuValAsnGlySerLeuHisIleSerSerSerLeuValProMetHisGlnGly 409
Db 1132 ACA----- 1134
Qy 410 SerLeuAlaAlaGlyPro---CysGlyCysSerSerCysLeuAsnIleGlySerLys 428
Db 1135 CCCACAGCTGGGGGCCACAGCTGGCTGCTGTG-TCGAGACTGCTGTGTGGAGCCAGG 1193
Qy 429 GlyLysSerSerTyrCysSerGluProThrLeuLeuLysLysAlaAsnLysLysGlnAsn 448
Db 1194 CTCAGAACTGCCCTCGCACCACCC-----AAGCTCTAGGCCCATGGACTGAAC 1244
Qy 449 PheLysSerGluValMetValProValLysAlaSerGlySer-----SerHisCys 465
Db 1245 ATCAGAGAGTCAGGGGTTTCGTCGCTGAGGGGTTGGAGGACGCTTCCCTCCTCACCCC 1304
Qy 466 ArgAlaGlyGln 469
Db 1305 AGAACTGGGGAG 1316
```

Search completed: March 15, 2003, 07:50:33  
Job time : 4293 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 13:47:15 ; Search time 7161 Seconds  
(without alignments)  
9944.779 Million cell updates/sec

Title: US-09-960-643-1  
Perfect score: 2447  
Sequence: 1 tggagtggagctcaagcag.....tttttcttaaaaaaaaaa 2447

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	2447	100.0	2447	6	AX399682	AX399682 Sequence
2	2445.4	99.9	2474	9	BC032787	BC032787 Homo sapi
3	2415	98.7	2464	9	AF428261	AF428261 Homo sapi
4	2390	97.7	2612	9	AK095713	AK095713 Homo sapi
5	1734.8	70.9	1738	9	HS2721161	AL049688 Human gen
6	1328.6	54.3	2416	10	BC021840	BC021840 Mus muscu
7	1328.2	54.3	2427	10	AF428262	AF428262 Mus muscu
8	905.6	37.0	157875	9	HS272116	AL023754 Human DNA
9	813	33.2	1013	10	D86557	D86557 Rattus norv
10	504.6	20.6	1579	9	AF286366	AF286366 Homo sapi
11	504.6	20.6	1671	6	AX167589	AX167589 Sequence
12	502.2	20.5	1074	6	AX166517	AX166517 Sequence
13	502.2	20.5	1074	6	AX167587	AX167587 Sequence
14	502.2	20.5	1158	6	AX167585	AX167585 Sequence
15	493	20.1	1439	10	RATCAMPKAA	L26288 Rattus norv
16	489.8	20.0	1480	9	HUMCKI	L41816 Homo sapien
17	485	19.8	1402	10	RATPRKI	L24907 Rattus norv
18	485	19.8	1448	10	BC014825	BC014825 Mus muscu
19	428.8	17.5	1618	10	AB004267	AB004267 Rattus no
20	428.4	17.5	1032	10	AB023027	AB023027 Mus muscu
21	428.4	17.5	1332	10	D86556	D86556 Rattus norv
22	428.4	17.5	1554	10	AF181984	AF181984 Mus muscu
23	427	17.4	1032	6	AX166520	AX166520 Sequence
24	418	17.1	1282	6	AX139102	AX139102 Sequence
25	291.4	11.9	1546	3	AB021864	AB021864 Caenorhab
26	290	11.9	3316	3	DMEI7917	Y17917 Drosophila
27	287.8	11.8	182054	10	AL365314	AL365314 Mouse DNA
28	287.8	11.8	184754	2	AC022675	AC022675 Mus muscu
29	250.4	10.2	3141	10	BC017634	BC017634 Mus muscu
30	247.6	10.1	191584	2	AC120477	AC120477 Rattus no
31	245.2	10.0	2897	6	AX099935	AX099935 Sequence
32	244.4	10.0	3655	10	RATCBVA	L22557 Rattus norv
33	243.6	10.0	1503	6	AX239863	AX239863 Sequence
34	243.6	10.0	2297	6	AX239861	AX239861 Sequence
35	243.6	10.0	3003	9	BC001921	BC001921 Homo sapi
36	243.6	10.0	3012	9	BC000497	BC000497 Homo sapi
37	243.6	10.0	3012	9	BC005828	BC005828 Homo sapi
38	243.6	10.0	3018	9	BC017363	BC017363 Homo sapi
39	243.6	10.0	3057	9	BC019256	BC019256 Homo sapi
40	243.6	10.0	3124	6	AX350342	AX350342 Sequence
41	229.8	9.4	42404	3	CBRG46J06	AC084656 Caenorhab
42	214.6	8.8	1701	8	EMECMKA	M74120 Emericella
43	197.2	8.1	3471	6	AR027895	AR027895 Sequence
44	194	7.9	3460	9	HS272212	AJ272212 Homo sapi
45	194	7.9	6612	12	AF416988	AF416988 Synthetic

ALIGNMENTS

RESULT 1  
AX399682  
LOCUS AX399682 2447 bp DNA linear PAT 06-JUN-2002  
DEFINITION Sequence 3 from Patent WO0224947.  
ACCESSION AX399682  
VERSION AX399682.1 GI:21335455  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Delaney,A.D. and Yoganathan,T.  
TITLE Cancer associated protein kinases and their uses  
JOURNAL Patent: WO 0224947-A 3 28-MAR-2002;

KINETEK PHARMACEUTICALS INC (CA); UNIV BRITISH COLUMBIA (CA)	
FEATURES	Location/Qualifiers
source	1..2447
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
CDS	70..1500
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAD33521.1"
	/db_xref="GI:21335456"
	/translation="MGRKEEDDCSWKKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRL TKGLKALCIKKSPAFRDSLENEIAVLKIKKHENIVTLEDIYESYTHYLVMOLVSG GELFDRIILGKVYTKDASLVIOQVLSAKVYLHENGIVHRDLKPENLLYLPOMENSKI NITDRLGLSKMEONGINSTACGTPGVYAVEVLAQRPYSKAVDCWSIGVTYILLGCGYP FYETESKLFERIKGDFEFPWDDISEAKDFICHLEKDPNERITCEKALUSHW IDGNLTALHRIQYVSLQIQKNFAKSKRWQAFNAAYVHHMRKLHMLNHSKLVPEVE NRPPTQASSETSRPSPSEITTEAPVLDHVALPALQPCQHGRRRTAPGGRSLNCL VNGSLHISSSLVPMHQGSLAAGPCSCSLNIGSKGSSYSCEPTLLIKRANKQNFK SEVMVPVKASGSSHCRAGQTGVCLIM"
BASE COUNT	590 a 707 c 604 g 546 t
ORIGIN	
Query Match 100.0%; Score 2447; DB 6; Length 2447;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 2447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 TGGAGTGGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCCCTCAGAACGCTTCAACT 60
Db	1 TGGNGTGGGAGCTCAGCAGGATCTTCCGAGTCCCTGGCATCCCTCAGAACGCTTCAACT 60
Qy	61 CTGGAGCAATGGGTGGAAGGAAGATGACTGCAGTTCCTTGGGAAGAAACAGACCACC 120
Db	61 CTGGAGCAATGGGTGGAAGGAAGATGACTGCAGTTCCTTGGGAAGAAACAGACCACC 120
Qy	121 AACATCCGGAACCTTCATTTTATGGAAGTCTCGGATCAGGAGCTTTCACAGAAGT 180
Db	121 AACATCCGGAACCTTCATTTTATGGAAGTCTCGGATCAGGAGCTTTCACAGAAGT 180
Qy	181 TTCTGTGTGAAGCAAGACTGACTGGGAAGCTCTTTTGCCTCAAGTGCATCAAGAAGTCA 240
Db	181 TTCTGTGTGAAGCAAGACTGACTGGGAAGCTCTTTTGCCTCAAGTGCATCAAGAAGTCA 240
Qy	241 CTTGCCCTCCGGACAGCAGCTGGAGATGAGATGAGATGCTGTGTGTAAGAAATCAAGCAT 300
Db	241 CTTGCCCTCCGGACAGCAGCTGGAGATGAGATGAGATGCTGTGTGTAAGAAATCAAGCAT 300
Qy	301 GAAACATTTGTACCTCGGAGGACATCTATGAGAGCACCACCACCTACTACTGGTGCATG 360
Db	301 GAAACATTTGTACCTCGGAGGACATCTATGAGAGCACCACCACCTACTACTGGTGCATG 360
Qy	361 CAGCTTCTTCTGGTGGGAGCTCTTTTACCCGGATCCTGGAGGGGGTGTCTACACAGAG 420
Db	361 CAGCTTCTTCTGGTGGGAGCTCTTTTACCCGGATCCTGGAGGGGGTGTCTACACAGAG 420
Qy	421 AAGGATCCAGTCTGGTATCCAGAGGCTCTTGGGAGTGAATACCTACATGAGAAAT 480
Db	421 AAGGATCCAGTCTGGTATCCAGAGGCTCTTGGGAGTGAATACCTACATGAGAAAT 480
Qy	481 GGCAATCTCCACAGAGACTTTAAAGCCCGGAAACCTGCTTTACCTTACCCTTGAAGAGAAC 540
Db	481 GGCAATCTCCACAGAGACTTTAAAGCCCGGAAACCTGCTTTACCTTACCCTTGAAGAGAAC 540
Qy	541 TCTAAGATCATGATCACTGACTTTGGTCTGTCCAAAGATGGAACAGAGATGGCATCATGTCC 600
Db	541 TCTAAGATCATGATCACTGACTTTGGTCTGTCCAAAGATGGAACAGAGATGGCATCATGTCC 600
Qy	601 ACTCCCTGTGGAGCCCGAGGCTACGTGGCTCCAGAGAGTGTGCGCCAGAGAACCCCTACAGC 660
Db	601 ACTCCCTGTGGAGCCCGAGGCTACGTGGCTCCAGAGAGTGTGCGCCAGAGAACCCCTACAGC 660
Qy	661 AAGSCTGTGGATTGCTGGTCCATCGGCGTCATCACCTACATATTTGCTCTGTGGATACCCCG 720
Db	661 AAGSCTGTGGATTGCTGGTCCATCGGCGTCATCACCTACATATTTGCTCTGTGGATACCCCG 720

Qy	721 CCGTTCTATGAAGAAACGGAGCTTAAGCTTTTCGAGAAGATCAAGGAGGCGCTACTATGAG 780
Db	721 CCGTTCTATGAAGAAACGGAGCTTAAGCTTTTCGAGAAGATCAAGGAGGCGCTACTATGAG 780
Qy	781 TTTGAGTCTCCATTTCTGGATGACATTTCTGAGTCAGCAAGGACTTTATTTGCCACTTG 840
Db	781 TTTGAGTCTCCATTTCTGGATGACATTTCTGAGTCAGCAAGGACTTTATTTGCCACTTG 840
Qy	841 CTTGAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGGCTTGTAGTCATCCCTGGATT 900
Db	841 CTTGAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGGCTTGTAGTCATCCCTGGATT 900
Qy	901 GACGGAACAGCGCCCTCCACCGGAGACATCTACCCATCAGTCAGCTCCAGATCCAGAAG 960
Db	901 GACGGAACAGCGCCCTCCACCGGAGACATCTACCCATCAGTCAGCTCCAGATCCAGAAG 960
Qy	961 AACTTTGCTAAGAGCAAGTGGAGCAAGCCCTTCAACAGCAGCAGCTGTGTGCACACATG 1020
Db	961 AACTTTGCTAAGAGCAAGTGGAGCAAGCCCTTCAACAGCAGCAGCTGTGTGCACACATG 1020
Qy	1021 AGGAAGCTACACATGAACCTGCACAGCCCGGCGTCCGCCAGAGTGGAGAACAGCGCG 1080
Db	1021 AGGAAGCTACACATGAACCTGCACAGCCCGGCGTCCGCCAGAGTGGAGAACAGCGCG 1080
Qy	1081 CCTGAAACTCAAGCTCAGAAACCTCTAGACCCAGCTCCCTGTGAGATCACCATCACCGAG 1140
Db	1081 CCTGAAACTCAAGCTCAGAAACCTCTAGACCCAGCTCCCTGTGAGATCACCATCACCGAG 1140
Qy	1141 GCACCTGTCTGGACCACAGTGTAGCACTCCCTGCCCTGACCCAAATTTACCTGCCAGCAT 1200
Db	1141 GCACCTGTCTGGACCACAGTGTAGCACTCCCTGCCCTGACCCAAATTTACCTGCCAGCAT 1200
Qy	1201 GGCGCGCGGCCCACTGGCCCTTGGTGGCAGTCCCTCAACTGCCCTGGTCAATGGCTCCCTC 1260
Db	1201 GGCGCGCGGCCCACTGGCCCTTGGTGGCAGTCCCTCAACTGCCCTGGTCAATGGCTCCCTC 1260
Qy	1261 CACATCAGCAGCAGCTGGTGCCCATGCAATCAGGGGTCCCTTGGCCGCCGGGCGCTGTGGC 1320
Db	1261 CACATCAGCAGCAGCTGGTGCCCATGCAATCAGGGGTCCCTTGGCCGCCGGGCGCTGTGGC 1320
Qy	1321 TGCTGCTCCAGCTGCTGAACATTTGGGAGCAAGGAAAGTCTCTACTGCTGTGAGCC 1380
Db	1321 TGCTGCTCCAGCTGCTGAACATTTGGGAGCAAGGAAAGTCTCTACTGCTGTGAGCC 1380
Qy	1381 ACACCTCTCAAAAAGGCCCAACAAAAACAGAACTTCAAGTCGGAGGTCATGGTACCAGTT 1440
Db	1381 ACACCTCTCAAAAAGGCCCAACAAAAACAGAACTTCAAGTCGGAGGTCATGGTACCAGTT 1440
Qy	1441 AAAGCCAGTGGCAGCTCCCACTGCGGCGAGGCGAGCTGGAGTCTGTCTCATTTATGTGA 1500
Db	1441 AAAGCCAGTGGCAGCTCCCACTGCGGCGAGGCGAGCTGGAGTCTGTCTCATTTATGTGA 1500
Qy	1501 TTCTCTGGAGCTGTGCTTATGTCATGCAATTTTTCAGGAGACATATTCAACTCCTCTGCT 1560
Db	1501 TTCTCTGGAGCTGTGCTTATGTCATGCAATTTTTCAGGAGACATATTCAACTCCTCTGCT 1560
Qy	1561 CTTCCAAACCTGTGTCTATCCGGCAGAGGAGGAGCAAGCAAGTGGAGAGGCGTT 1620
Db	1561 CTTCCAAACCTGTGTCTATCCGGCAGAGGAGGAGCAAGCAAGTGGAGAGGCGTT 1620
Qy	1621 AGCAGGAGCAGTTTCTGGCCAGAGACACAGCTGCTGCCAGCGGGGAGCCCTCATAG 1680
Db	1621 AGCAGGAGCAGTTTCTGGCCAGAGACACAGCTGCTGCCAGCGGGGAGCCCTCATAG 1680
Qy	1681 GAGGCCCCAGGAGGAGCCCCAAGCGGTAGAAGCCCTTGTGAAGCTGTGAGCAGGAGAAGC 1740
Db	1681 GAGGCCCCAGGAGGAGCCCCAAGCGGTAGAAGCCCTTGTGAAGCTGTGAGCAGGAGAAGC 1740
Qy	1741 GGTGCCACACAGCTTCCAGGTCTCCCTGACCTGCTCTCTTGTGCCACACACCTTACGTG 1800
Db	1741 GGTGCCACACAGCTTCCAGGTCTCCCTGACCTGCTCTCTTGTGCCACACACCTTACGTG 1800

QY	1801	CCGTGGCTGTGTCAGTGTACGTAGATAGCTCTCCCTGGGTCGTGCTGTTTGTGCGA	1860
Db	1801		
QY	1861	AAAGCTTAATGGCTGGCCAGCTGTGTCACCTTCTCCAAGCAAAAGCCATATGAGGCATC	1920
Db	1861		
QY	1921	TACCAGAGCTCCACTCTGCACACTCAGTCCAGCTCTCAAGCTTCCACCTCTTGGC	1980
Db	1921		
QY	1981	CAGATTGGCTCATTAATGCTTTCCTGCCCTCCCATTCGATGAATGACAGCAGCTCCCA	2040
Db	1981		
QY	2041	TGGTGTGCTGCTGTGAGCTCTTCAAGTTCCTTAATCCTTAACCTCCAGGATTAGCTCCCAAG	2100
Db	2041		
QY	2101	TGCGCTGAGACCCAGCCAGCACATCTCGGCCCTTCTCCCTGCCTCAATCTAAAGCAGT	2160
Db	2101		
QY	2161	GCCACACCTCCAAAGTGGATAGAAAGTTCATGAGTAAAGGCTGCAAGGAATCTT	2220
Db	2161		
QY	2221	ATCCTGGCCACATGCTCTCGGTGCACACACCAATGAGGATTAACTTGGAAAGTTGACTAT	2280
Db	2221		
QY	2281	TTTAATGCTGCGAGAGTCTTAATCCTCGCTCTGTTCCCTTTCTCTCTTGAAGTCC	2340
Db	2281		
QY	2341	AGCACACCATTTCTGCTTCCCTCCAGTTTCTCGCCCTCCACCCCTCCAGCTTTCATGCTC	2400
Db	2341		
QY	2401	AGTGTTGTGCTTAATAAATGACATATTTTCTCTTAAAAAAAAA	2447
Db	2401		
RESULT 2	BC032787		
LOCUS	BC032787	2474 bp	linear
DEFINITION	Homo sapiens, calcium/calmodulin-dependent protein kinase IG, clone PRI 27-JUN-2002		
ACCESSION	MGC:44894 IMAGE:5179957, mRNA, complete cds.		
VERSION	BC032787		
KEYWORDS	MGC.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	1 (bases 1 to 2474)		
TITLE	Strausberg, R.		
Submitted (06-JUN-2002)	National Institutes of Health, Mammalian		
Gene Collection (MGC),	Cancer Genomics Office, National Cancer		
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,	USA		
NIH-MGC Project URL:	http://mgc.nci.nih.gov		
Contact: MGC help desk			
Email: cgapbs-r@mail.nih.gov			
Tissue Procurement:	Life Technologies, Inc.		
cDNA Library Preparation:	Life Technologies, Inc.		
DNA Sequencing by:	The I.M.A.G.E. Consortium (LLNL)		
Sequencing Center (NISC),	Gaithersburg, Maryland,		

Web site: http://www.nisc.nih.gov/ Contact: nisc.mgc@nih.gov	
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgoev, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov	
Series: IRAK Plate: 68 Row: k Column: 17	
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 14196444.	
FEATURES	Location/Qualifiers
source	1..2474
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/db_xref="taxon:57172"
	/clone="MGC:44894 IMAGE:5179957"
	/tissue.type="Brain, Lung, Testis, adult, pooled whole"
	/clone_lib="NIH_MGC_115"
	/lab_host="DH10B"
	/note="Vector: PCMV-SPORT6"
	/size="1500"
CDS	70..1500
	/codon_start=1
	/product="calcium/calmodulin-dependent protein kinase IG"
	/protein_id="AAH32787.1"
	/db_xref="GI:21619665"
	/translation="MGRKEDDCSSKKQTNIRKTFIFMEVILSGAFSEFLVKQRL TKGLFALKIKSPAFRDSLENEIAVLKIKHENIVTLEDIESTHYILVQLVSK GEFLRILRGVYTKDASLVQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKI MTDFGLKMEQNGINSTACPTGYVAPEVLAQKPYSKAVDCWSTGVTYTLGCGYP FYETEESKLFKEIKRGYVEFFSDFWDDISAKDFICHLLEKDPNERTYCKALSHPW IDGNTALHDIYPSVLOIKNFASKRWAFNAVAHVHMKLHMLNHSVGRPEVE NRPPQASSTSRPSSPRITTEAPVLDHVALPALTOLPCQHGRRPTAPGGRSLNCL VNGSLHSSSLVPMHOGSLAAGPCGCCSCLNIGSKGSSYCSSEPTLLKANKKQNFK SEVMVPKASSSHCRAGOTGVCLIM"
BASE COUNT	618 a 707 c 603 g 546 t
ORIGIN	
Query Match 99.9%; Score 2445.4; DB 9; Length 2474;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 2446; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 TGGAGTGGGAGCTCAAGCAGGATTCCTCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT 60
Db	
QY	1 TGGAGTGGGAGCTCAAGCAGGATTCCTCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT 60
Db	
QY	61 CTGGAGGCAATGGGTGCAAGGAGAGATGACTGCAGTTCTCTGGAGAAACAGACCACC 120
Db	
QY	61 CTGGAGGCAATGGGTGCAAGGAGAGATGACTGCAGTTCTCTGGAGAAACAGACCACC 120
Db	
QY	121 AACATCCGGAACACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCAGAAAGTT 180
Db	
QY	121 AACATCCGGAACACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCAGAAAGTT 180
Db	
QY	181 TTCTCTGGTGAAGCAAGACATGACTGGGAAGCTCTTTGCTCTCAAGTGCATCAAGAAGTCA 240
Db	
QY	181 TTCTCTGGTGAAGCAAGACATGACTGGGAAGCTCTTTGCTCTCAAGTGCATCAAGAAGTCA 240
Db	
QY	241 CTGTCCTCTCCGGACAGCAGCTGGAGATGAGATTGCTGTCTTGAAGAAGATCAAGCAT 300
Db	
QY	241 CTGTCCTCTCCGGACAGCAGCTGGAGATGAGATTGCTGTCTTGAAGAAGATCAAGCAT 300
Db	
QY	301 GAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCACCACCTACTACTGCTGTCATG 360
Db	
QY	301 GAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCACCACCTACTACTGCTGTCATG 360
Db	
QY	361 CAGCTTGTCTTGGTGGGAGCTCTTTGACCGGATCTCTGGAGGGGGTGTCTTACACAGAG 420
Db	

Db 361 CAGCTTGTTCTGGTGGGAGCTCTTTGACCGGATCTTGAGCGGGGTGCTACACAGAG 420  
QY 421 AAGGATGCCAGTCTGGTGATCCAGCAGGTCTTTGTCGGCAGTGAATACTACATGAGAAT 480  
Db 421 AAGGATGCCAGTCTGGTGATCCAGCAGGTCTTTGTCGGCAGTGAATACTACATGAGAAT 480  
QY 481 GGCATGCTCCACAGAGACTTAAAGCCGAAACCTGCTTTTACCTTTACCCCTGAAGAGAAC 540  
Db 481 GGCATGCTCCACAGAGACTTAAAGCCGAAACCTGCTTTTACCTTTACCCCTGAAGAGAAC 540  
QY 541 TCTAAGATCATGATCACTGACTTTGGTCTGTGTCCTCAAGATGGAACAGAAATGGCATATGAG 600  
Db 541 TCTAAGATCATGATCACTGACTTTGGTCTGTGTCCTCAAGATGGAACAGAAATGGCATATGAG 600  
QY 601 ACTGCGTGTGGACCCAGGCTACCTGGCTCCAGAAAGTCTGGCCAGAAACCTTACAGC 660  
Db 601 ACTGCGTGTGGACCCAGGCTACCTGGCTCCAGAAAGTCTGGCCAGAAACCTTACAGC 660  
QY 661 AAGGCTGTGGATTTGCTGGTCCATCGCGCTCATCACTTACATATTTGCTGTGGATACCC 720  
Db 661 AAGGCTGTGGATTTGCTGGTCCATCGCGCTCATCACTTACATATTTGCTGTGGATACCC 720  
QY 721 CCGTCTCTAGAGAAACGAGTCTAAGCTTTTCGAGAGATCAAGAGGGCTACTATGAG 780  
Db 721 CCGTCTCTAGAGAAACGAGTCTAAGCTTTTCGAGAGATCAAGAGGGCTACTATGAG 780  
QY 781 TTTGAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGCAAGGACTTTATTTGCCACTTG 840  
Db 781 TTTGAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGCAAGGACTTTATTTGCCACTTG 840  
QY 841 CTTGAGAAGATCCGAACGAGCGGTACACCTGTGTGAGAGGCTTTGAGTCATCCCTGGATT 900  
Db 841 CTTGAGAAGATCCGAACGAGCGGTACACCTGTGTGAGAGGCTTTGAGTCATCCCTGGATT 900  
QY 901 GAGGGAACACAGCCCTCCACCGGACATCTACCCATCAGTCAGCTCCAGATCCAGAG 960  
Db 901 GAGGGAACACAGCCCTCCACCGGACATCTACCCATCAGTCAGCTCCAGATCCAGAG 960  
QY 961 AACTTTGCTAAGAGCAAGTGGAGCAAGCTTTCAACGAGCAGCTGTGTGTCACCAATG 1020  
Db 961 AACTTTGCTAAGAGCAAGTGGAGCAAGCTTTCAACGAGCAGCTGTGTGTCACCAATG 1020  
QY 1021 AGGAAGCTACATGAACTTGCACAGCCGGCGGTCCGCGGAGAGGTTGGAGAACAGCGCG 1080  
Db 1021 AGGAAGCTACATGAACTTGCACAGCCGGCGGTCCGCGGAGAGGTTGGAGAACAGCGCG 1080  
QY 1081 CTTGAAACTCAAGCCCTCAGAAACCTTAGAACAGCTCCCTTGAGATCACCATCACCAG 1140  
Db 1081 CTTGAAACTCAAGCCCTCAGAAACCTTAGAACAGCTCCCTTGAGATCACCATCACCAG 1140  
QY 1141 GCACCTGTCTTGACACAGTGTAGACCTCCCTGAGATCCCTGACCCAAATTTACCTTGCCAGAT 1200  
Db 1141 GCACCTGTCTTGACACAGTGTAGACCTCCCTGAGATCCCTGACCCAAATTTACCTTGCCAGAT 1200  
QY 1201 GGGCGCGGCCACTGCCCTGGTGGCAGGTCCCTCACTGCTGCTGCTCAATGGCTCCCTC 1260  
Db 1201 GGGCGCGGCCACTGCCCTGGTGGCAGGTCCCTCACTGCTGCTGCTCAATGGCTCCCTC 1260  
QY 1261 CACATCAGCAGAGCTGTGTGCTCAGTATCAGGGGTCCCTGGCGCGGGCCCTGTGGC 1320  
Db 1261 CACATCAGCAGAGCTGTGTGCTCAGTATCAGGGGTCCCTGGCGCGGGCCCTGTGGC 1320  
QY 1321 TGCTGTCTCCAGCTGCTGAACTTTGGAGCAAGGAAAGTCTCTCTACTGCTCTGAGGCC 1380  
Db 1321 TGCTGTCTCCAGCTGCTGAACTTTGGAGCAAGGAAAGTCTCTCTACTGCTCTGAGGCC 1380  
QY 1381 ACACCTCTCAAAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTATGATGACAGTT 1440  
Db 1381 ACACCTCTCAAAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTATGATGACAGTT 1440  
QY 1441 AAAGCCAGTGGCAGCTCCCACTGCCGGCAGGCGAGACTGGAGTCTGTCTATTATGGA 1500  
Db 1441 AAAGCCAGTGGCAGCTCCCACTGCCGGCAGGCGAGACTGGAGTCTGTCTATTATGGA 1500

QY 1501 TTCTTGAGCCTCTGCTATGTCTACCTGCAATTTTCAGGAGACATATTTCAACTCCTCTGCT 1560  
Db 1501 TTCTTGAGCCTCTGCTATGTCTACCTGCAATTTTCAGGAGACATATTTCAACTCCTCTGCT 1560  
QY 1561 CTTTCCAAACCTGTGTCTATCCGCGCAGAGGGAGGAAGGAGAGCAAGTGGAGCAGGCTT 1620  
Db 1561 CTTTCCAAACCTGTGTCTATCCGCGCAGAGGGAGGAAGGAGAGCAAGTGGAGCAGGCTT 1620  
QY 1621 AGCAGGAGCAGTTTCTGGCCAGAGAACACAGCCTGTGCCAGGGGGCAGCCCTCATAG 1680  
Db 1621 AGCAGGAGCAGTTTCTGGCCAGAGAACACAGCCTGTGCCAGGGGGCAGCCCTCATAG 1680  
QY 1681 GAGGCCAGGAGGAGCCCAAGCGGTAGAAAGCTTTTGAAGCTGTGAGCAGGAGAAAG 1740  
Db 1681 GAGGCCAGGAGGAGCCCAAGCGGTAGAAAGCTTTTGAAGCTGTGAGCAGGAGAAAG 1740  
QY 1741 GGTGCCACACAGTTCAGGTCTCCCTGACCTGCTGCTGCTATGCCCCACACCTTACGTG 1800  
Db 1741 GGTGCCACACAGTTCAGGTCTCCCTGACCTGCTGCTGCTATGCCCCACACCTTACGTG 1800  
QY 1801 CCGTGGCTCTGTCAGTGTACGTAGATAGCTCTGCTGCTGGGTCTGTGTTGTCTGTA 1860  
Db 1801 CCGTGGCTCTGTCAGTGTACGTAGATAGCTCTGCTGCTGGGTCTGTGTTGTCTGTA 1860  
QY 1861 AAAGCTTAATGGCTGGCCAGGCTGTGTACCTTTCTCCAAAGCAAAAGCCATATATGGAGCATC 1920  
Db 1861 AAAGCTTAATGGCTGGCCAGGCTGTGTACCTTTCTCCAAAGCAAAAGCCATATATGGAGCATC 1920  
QY 1921 TACCCAGACTCCCACTCTGCACACACTCACTCCCACTCTCAAGCCTTCAACCTCTTGGC 1980  
Db 1921 TACCCAGACTCCCACTCTGCACACACTCACTCCCACTCTCAAGCCTTCAACCTCTTGGC 1980  
QY 1981 CAGATTGGGCTCATTAATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
Db 1981 CAGATTGGGCTCATTAATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
QY 2041 TGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
Db 2041 TGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
QY 2101 TGGCTCAGACCCAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2160  
Db 2101 TGGCTCAGACCCAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2160  
QY 2161 GCCACACCTCCAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2220  
Db 2161 GCCACACCTCCAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2220  
QY 2221 ATCTGCGCCACATGCTCTCCGTCGACACACACACACACACACACACACACACACAC 2280  
Db 2221 ATCTGCGCCACATGCTCTCCGTCGACACACACACACACACACACACACACACACAC 2280  
QY 2281 TTTAATGTCTGCCAGGAGTTCTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340  
Db 2281 TTTAATGTCTGCCAGGAGTTCTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340  
QY 2341 AGCAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400  
Db 2341 AGCAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400  
QY 2401 AGTGTGCTGCTTAATAAATGGACATATTTTCTCTAAAAA 2447  
Db 2401 AGTGTGCTGCTTAATAAATGGACATATTTTCTCTAAAAA 2447

RESULT 3  
AF428261  
LOCUS  
DEFINITION  
ACCESSION  
VERSION

AF428261  
Homo sapiens calcium/calmodulin-dependent protein kinase I gamma  
(CAMK1G) mRNA, complete cds.  
AF428261  
AF428261.1 GI:16755791

2464 bp mRNA linear PRI 07-NOV-2001

KEYWORDS Homo sapiens.  
SOURCE Homo sapiens  
ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 2464)  
AUTHORS Schutte,B.C., Bjork,B.C., Coppage,K.B., Malik,M.I., Gregory,S.G.,  
Scott,D.J., Brenzelle,L.M., Watanabe,Y., Dixon,M.J. and Murray,J.C.  
TITLE A preliminary gene map for the Van der Woude syndrome critical  
region derived from 900 kb of genomic sequence at 1q32-q41  
JOURNAL Genome Res. 10 (1), 81-94 (2000)  
MEDLINE 20113118  
PubMed 10645953  
REFERENCE 2 (bases 1 to 2464)  
AUTHORS Bjork,B.C., Watanabe,Y., Murray,J.C. and Schutte,B.C.  
TITLE Characterization of the human ortholog of rat Cam Kinase I gamma  
(CamK1g) at 1q32-q41  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 2464)  
AUTHORS Bjork,B.C., Watanabe,Y., Murray,J.C. and Schutte,B.C.  
TITLE Direct Submission  
JOURNAL Submitted (05-OCT-2001) Genetics, University of Iowa, 140 EMRB,  
Iowa City, IA 52242, USA  
FEATURES  
source Location/Qualifiers  
1..2464  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="1q32-q41"  
1..2464  
/gene="CAMK1g"  
73..1503  
/gene="CAMK1g"  
/codon\_start=1  
/product="calcium/calmodulin-dependent protein kinase I  
gamma"  
/protein\_id="AAL28100.1"  
/db\_xref="GI:16755792"  
/translation="MGRKEEDCSWKKQTINIKTFIFMEVLGSAFSEVFLVKQRL  
TKLPAKICIKSPAFROSLLENEIAVLKIKHENIVLEDIYESTHIYLVMLQVSG  
GEILFRLRMEGVYTERDASLVQQVLSAVKYLHENGIVHRDLKPNLDLYLPENSKI  
MITDFLSRNGINGIMSTAGTGPYVAPEVLAQPKYSKAVDCWSIGVITYILLGCGPP  
FYETEKSLEFIKEGYEFESFPWDDISEAKDFICHLLEKDPNERTYCEKALSHPW  
IDGNTALHRDIYVPSLSIQKFNFAKSKWQAFNAAVVHMRKLHMLHSPGIRPEVE  
NRPEPTQASSTRSPSPETITTEAPVLDHVALPALPOLPCQHGRRPTAPGGRSINCL  
VNGSLHISSLVPMHQGSLAAGPCGCCSCLNIGNKGSYSYCSPTLLKANKQKPK  
SEVMVPKASGSHCRAGTGVCLIN"  
polyA\_signal 2415..2420  
BASE COUNT 609 a 708 c 602 g 545 t  
ORIGIN  
Query Match 98.7%; Score 2415; DB 9; Length 2464;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2418; Conservative 0; Mismatches 5; Indels 1; Gaps 1:  
Qy 24 TCTTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACTCTGGAGCAATGGGTCGAAGA 83  
Db 27 TCTTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACTCTGGAGCAATGGGTCGAAGA 86  
Qy 84 AGAAGATGACTGAGTTCCTGGAGAAACAGACACACACATCCGGAAAACCTTCATTTT 143  
Db 87 AGAAGATGACTGAGTTCCTGGAGAAACAGACACACATCCGGAAAACCTTCATTTT 146  
Qy 144 TATGAAGTGTGGATCAGAGCTTCTCAGAAGTTTCTCGTGGTGAAGCAAGACTGAC 203  
Db 147 TATGAAGTGTGGATCAGAGCTTCTCAGAAGTTTCTCGTGGTGAAGCAAGACTGAC 206  
Qy 204 TGGGAAGCTCTTGTCTGTAAGTGCATCAAGAAGTCACTGCTTCCGGACAGCAGCCT 263  
Db 207 TGGGAAGCTCTTGTCTGTAAGTGCATCAAGAAGTCACTGCTTCCGGACAGCAGCCT 266  
Qy 264 GGAGATGAGATTGCTGTGTGAAAAGATCAAGCATGAAACATTGTGACCTGGAGGA 323

Db 267 GGAGATGAGATTGCTGTGTGAAAAGATCAAGCATGAAAACATTGTGACCTGGAGGA 326  
Qy 324 CATCTATGAGAGACACACACACTACTACCTGGTTCATGCAGCTTGTTCCTGGTGGGAGCT 383  
Db 327 CATCTATGAGAGACACACACACTACTACCTGGTTCATGCAGCTTGTTCCTGGTGGGAGCT 386  
Qy 384 CTTTGACCGGATCTGAGAGGGGTGTCTACACAGAGAGAGTGCAGCTGTGGTGATCCA 443  
Db 387 CTTTGACCGGATCTGAGAGGGGTGTCTACACAGAGAGAGTGCAGCTGTGGTGATCCA 446  
Qy 444 GCAGGTCTTGTGGCAGTGAATACCTACATGAGATGGCATGCTCCACAGAGACTTAA 503  
Db 447 GCAGGTCTTGTGGCAGTGAATACCTACATGAGATGGCATGCTCCACAGAGACTTAA 506  
Qy 504 GCCCGAAAACCTGCTTTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCACTGACTT 563  
Db 507 GCCCGAAAACCTGCTTTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCACTGACTT 566  
Qy 564 TGGTCTGTCCAAGATGAACAGATGGCATGCTTCCACTGCTGTGGACCCAGGCTA 623  
Db 567 TGGTCTGTCCAAGATGAACAGATGGCATGCTTCCACTGCTGTGGACCCAGGCTA 626  
Qy 624 CGTGGCTCCAGAGTGTGCGCCAGAAACCTTACAGCAAGGCTGTGGATTGTGTGCTCAT 683  
Db 627 CGTGGCTCCAGAGTGTGCGCCAGAAACCTTACAGCAAGGCTGTGGATTGTGTGCTCAT 686  
Qy 684 CGGCGTCATCACTTACATATTGCTGTGGATACCCCGCTTCTATGAAGAAACGAGTC 743  
Db 687 CGGCGTCATCACTTACATATTGCTGTGGATACCCCGCTTCTATGAAGAAACGAGTC 746  
Qy 744 TAAGCTTTTCGAGAAGATCAAGAGGCTACTATGATTGAGTCTCCATCTGGGATGA 803  
Db 747 TAAGCTTTTCGAGAAGATCAAGAGGCTACTATGATTGAGTCTCCATCTGGGATGA 806  
Qy 804 CATTTCTGAGTCAGCCAAAGACTTTATTGCCCCTTGTGTGAGAAGATCCGAACGAGCG 863  
Db 807 CATTTCTGAGTCAGCCAAAGACTTTATTGCCCCTTGTGTGAGAAGATCCGAACGAGCG 866  
Qy 864 GTACACCTGTGAGAAGGCTTTGAGTCACTCCTGGATTGAGCGAAACACAGCCCTCCACCG 923  
Db 867 GTACACCTGTGAGAAGGCTTTGAGTCACTCCTGGATTGAGCGAAACACAGCCCTCCACCG 926  
Qy 924 GGACATCTACCCATCAGTCAGCTCCAGCTCCAGATCCAGAGAACTTTGCTAAAGCAAGTGGAG 983  
Db 927 GGACATCTACCCATCAGTCAGCTCCAGCTCCAGATCCAGAGAACTTTGCTAAAGCAAGTGGAG 986  
Qy 984 GCAAGCCTTCAACGACGAGCTGTGGTGCACCATGAGGAACTTACATGAACCTGCA 1043  
Db 987 GCAAGCCTTCAACGACGAGCTGTGGTGCACCATGAGGAACTTACATGAACCTGCA 1046  
Qy 1044 CAGCCCGGCGTCCGCCAGAGGTGGAGAACAGCGCCCTGAAACTCAAGCCTCAGAAAC 1103  
Db 1047 CAGCCCGGCGTCCGCCAGAGGTGGAGAACAGCGCCCTGAAACTCAAGCCTCAGAAAC 1106  
Qy 1104 CTCTAGACCCAGCTCCCTCTGAGATCACCATCACGAGGACCTGTCTCTGGACCAAGTGT 1163  
Db 1107 CTCTAGACCCAGCTCCCTCTGAGATCACCATCACGAGGACCTGTCTCTGGACCAAGTGT 1166  
Qy 1164 AGCACTCCCTGCCCTGACCCAAATTTACCTGCCAGATGGCGCCGCGCCACTGCCCTGG 1223  
Db 1167 AGCACTCCCTGCCCTGACCCAAATTTACCTGCCAGATGGCGCCGCGCCACTGCCCTGG 1226  
Qy 1224 TGGCAGGTCCCTCAACTGCTTGGTCAATGGCTCCCTCCATCAGCAGCAGCCTGGTGCC 1283  
Db 1227 TGGCAGGTCCCTCAACTGCTTGGTCAATGGCTCCCTCCATCAGCAGCAGCCTGGTGCC 1286  
Qy 1284 CATGCATCAGGGGTCCCTTGGCGCCGGGCGCTGTGGCTGTCTGTCCAGCTGCCTGACAT 1343  
Db 1287 CATGCATCAGGGGTCCCTTGGCGCCGGGCGCTGTGGCTGTCTGTCCAGCTGCCTGACAT 1346  
Qy 1344 TGGGAGCAAGGAAGTTCCTCTACTGCTGTGAGCCACACTCCTCAAAAAGGCCAACAA 1403

```
Db 1347 TGGGAACAAAAGGAAGTCCCTCTACTGCTCTGAGCCCAACACCTCTCTCAAAAAGGCCAACAA 1406
Qy 1404 AAAACAGAACTTCAAGTCGAGGTCTATGGTACCAGTAAAGCCAGTGGCAGTCCCACTG 1463
Db 1407 AAAACAGAACTTCAAGTCGAGGTCTATGGTACCAGTAAAGCCAGTGGCAGTCCCACTG 1466
Qy 1464 CCGGGCAGGCGAGACTGGAGTCTGTCTCTCAATATGTATGTCTCTGGAGCCCTGTGCTATGTC 1523
Db 1467 CCGGGCAGGCGAGACTGGAGTCTGTCTCTCAATATGTATGTCTCTGGAGCCCTGTGCTATGTC 1526
Qy 1524 ACTGCAATTTTCAGGAGACATATTAACACTCTCTGCTCTTCCAAACCTGGTGTCTATCG 1583
Db 1527 ACTGCAATTTTCAGGAGACATATTAACACTCTCTGCTCTTCCAAACCTGGTGTCTATCG 1586
Qy 1584 GCAGAGGAGGAGGCGAGCAAGTGGAGCAGGCTTAGCAGAGCAGTTCCTGGCCAGA 1643
Db 1587 GCAGAGGAGGAGGCGAGCAAGTGGAGCAGGCTTAGCAGAGCAGTTCCTGGCCAGA 1646
Qy 1644 AGCACAGCCTGTGTCAGCGGGGCGAGCCCTCATAGGAGGCCAGGAGGAGGCCCAAG 1703
Db 1647 AGCACAGCCTGTGTCAGCGGGGCGAGCCCTCATAGGAGGCCAGGAGGAGGCCCAAG 1706
Qy 1704 GCGTAGAAGCCTGTTGAAGCTGTGAGCAGAGAGAGGGGTGCCACACGCTCCAGGTCT 1763
Db 1707 GCGTAGAAGCCTGTTGAAGCTGTGAGCAGAGAGAGGGGTGCCACACGCTCCAGGTCT 1766
Qy 1764 CCCTGAGCTGCTCTATGCCCCACACCCCTACGTGCCGTGCTCTGTGCAGTGTACGT 1823
Db 1767 CCCTGAGCTGCTCTATGCCCCACA -CCTACGTGCCGTGCTCTGTGCAGTGTACGT 1825
Qy 1824 AGATAGCTCTGGCCTGGGTCTGTGCTGTTTGTGCTGAAAGCTTAATGGGCTGGCCAGCG 1883
Db 1826 AGATAGCTCTGGCCTGGGTCTGTGCTGTTTGTGCTGAAAGCTTAATGGGCTGGCCAGCG 1885
Qy 1884 TGTGTACCTTCTCCAGCAAGCAATATAGCAGCATCTACCAGACTCCCACTCTGCACA 1943
Db 1886 TGTGTACCTTCTCCAGCAAGCAATATAGCAGCATCTACCAGACTCCCACTCTGCACA 1945
Qy 1944 CACTCACTCCCACTCTCAAGCCTCCAACTCTTGGCCAGATTGGCTCATTAATGCTGT 2003
Db 1946 CACTCACTCCCACTCTCAAGCCTCCAACTCTTGGCCAGATTGGCTCATTAATGCTGT 2005
Qy 2004 TGCCTGCCCTCTGCATGAATGACAGGCAGCTCCCATGGTGTCTGCTCTGAGTGTCTTT 2063
Db 2006 TGCCTGCCCTCTGCATGAATGACAGGCAGCTCCCATGGTGTCTGCTCTGAGTGTCTTT 2065
Qy 2064 CAAGTCTTAATCTTAAGTCCAGGATTAGCTCCCAAGTGGCTGAGACCCAGCCAGCACA 2123
Db 2066 CAAGTCTTAATCTTAAGTCCAGGATTAGCTCCCAAGTGGCTGAGACCCAGCCAGCACA 2125
Qy 2124 CTTCGCGCCTTCTCCCTGCTCAATCTAAAGCAGTGGCCACACCCTCCAAAGTGAATA 2183
Db 2126 CTTCGCGCCTTCTCCCTGCTCAATCTAAAGCAGTGGCCACACCCTCCAAAGTGAATA 2185
Qy 2184 GAAAGAGTTCATGAGTAAAGGCTGCAAGCAATTTATCTGCGCCACATGCTCCCGG 2243
Db 2186 GAAAGAGTTCATGAGTAAAGGCTGCAAGCAATTTATCTTCTGCGCCACATGCTCCCGG 2245
Qy 2244 CACACACCAATGAGTAAACCTTGGAAAGTTGACTATTTTAAATGTCTGCCAGGAGTTCTA 2303
Db 2246 CACACACCAATGAGTAAACCTTGGAAAGTTGACTATTTTAAATGTCTGCCAGGAGTTCTA 2305
Qy 2304 ATCCTGCTCTGTTCCTTTCTCTCTGAAAGTCCAGCAGCAGCAATTTCTGTCTTCCC 2363
Db 2306 ATCCTGCTCTGTTCCTTTCTCTCTGAAAGTCCAGCAGCAGCAATTTCTGTCTTCCC 2365
Qy 2364 CAGTTTCTGCGCCTCCACCCCTCCAGCTTCATGCTCAGTGTGTGCTTAATAAATGGA 2423
Db 2366 CAGTTTCTGCGCCTCCACCCCTCCAGCTTCATGCTCAGTGTGTGCTTAATAAATGGA 2425
Qy 2424 CATATTTTCTCTAAAAAATA 2447
Db 2426 CATATTTTCTCTAAAAAATA 2449
```

AK095713 2612 bp mRNA linear PRI 15-JUL-2002  
Homo sapiens cDNA FLJ38394 fis, clone FEBRA2007534, moderately similar to CamK1-like protein kinase.

AK095713  
AK095713.1 GI:21755035  
oligo capping: fis (full insert sequence).  
Homo sapiens fetal brain cDNA to mRNA, clone\_lib:FEBRA2 clone:FEBRA2007534.

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1  
Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hirao,M., Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.  
NEDO human cDNA sequencing project

Unpublished  
2 (bases 1 to 2612)  
Isogai,T. and Yamamoto,J.  
Direct Submission  
Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES  
Location/Qualifiers  
1..2612  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="FEBRA2007534"  
/tissue\_type="brain"  
/clone\_lib="FEBRA2"  
/dev\_stage="fetal"  
/note="cloning vector: pME18SFL3"

BASE COUNT 619 a 750 c 645 g 598 t

ORIGIN  
Query Match 97.7%; Score 2390; DB 9; Length 2612;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2393; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 39 GCATCTCTCAGAGCTTCAACTCTGGAGCAATGGTTCGAAAGAGAGATGACTGCAG 98  
Db 215 GGCATCTCTCAGAGCTTCAACTCTGGAGCAATGGTTCGAAAGAGAGATGACTGCAG 274  
Qy 99 TTCCTGGAGAAACAGACACCACATCCGAAACCTTCATTTTATGGAAGTCTGTGG 158  
Db 275 TTCCTGGAGAAACAGACACCACATCCGAAACCTTCATTTTATGGAAGTCTGTGG 334  
Qy 159 ATCAGAGCTTTCCTCAGAAGTTTTCCTGGTGAAGCAAGACTGACTGGGAAGCTTTTCG 218  
Db 335 ATCAGAGCTTTCCTCAGAAGTTTTCCTGGTGAAGCAAGACTGACTGGGAAGCTTTTCG 394  
Qy 219 TCTGAAGTGCATCAGAGTCCACCTGCTCCGGACAGCAGCCCTGGAGATGAGATTGC 278  
Db 395 TCTGAAGTGCATCAGAGTCCACCTGCTCCGGACAGCAGCCCTGGAGATGAGATTGC 454  
Qy 279 TGTGTTGAAAAGATCAAGCATGAAACATTTGTGACCCCTGGAGGACATCTATGAGACAC 338

|||||  
Db 455 TGTGTTGAAAAAGATCAAGCATGAAAAAATTTGTGACCTGGAGGACATCTATGAGAGCAC 514  
QY 339 CACCCACTACTACCTGGTCATGACGCTTGTTCTGTTGGTGGGAGCTCTTTGACCCGATCCT 398  
Db 515 CACCCACTACTACCTGGTCATGACGCTTGTTCTGTTGGTGGGAGCTCTTTGACCCGATCCT 574  
QY 399 GGAGCGGGTGTCTACACAGAGAAGGATGCCAGTGTGCTGATGCCAGAGGCTCTTGTGGGC 458  
Db 575 GGAGCGGGTGTCTACACAGAGAAGGATGCCAGTGTGCTGATGCCAGAGGCTCTTGTGGGC 634  
QY 459 AGTGAATACCTACATGAGATGGCATGCTCCACAGAGACTTAAAGCCGAAAACCTTGCT 518  
Db 635 AGTGAATACCTACATGAGATGGCATGCTCCACAGAGACTTAAAGCCGAAAACCTTGCT 694  
QY 519 TTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCACTGACTTTGCTGTGCTCAAGAT 578  
Db 695 TTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCACTGACTTTGCTGTGCTCAAGAT 754  
QY 579 GGAACAGATGGCATGCTGCACCTGCCTGTGGGACCCAGGCTACGTGGCTCCAGAGT 638  
Db 755 GGAACAGATGGCATGCTGCACCTGCCTGTGGGACCCAGGCTACGTGGCTCCAGAGT 814  
QY 639 GCTGGCCAGAAACCTCAGCAAGGCTGTGGATTGCTGGTCCATCGGCGTCATCACTA 698  
Db 815 GCTGGCCAGAAACCTCAGCAAGGCTGTGGATTGCTGGTCCATCGGCGTCATCACTA 874  
QY 699 CATATTGCTCTGTGGATACCCGCCCTCTATGAGAAACGGAGTCTAAGCTTTTCGAGAA 758  
Db 875 CATATTGCTCTGTGGATACCCGCCCTCTATGAGAAACGGAGTCTAAGCTTTTCGAGAA 934  
QY 759 GATCAAGGAGGCTACTATGAGTTTGTAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGC 818  
Db 935 GATCAAGGAGGCTACTATGAGTTTGTAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGC 994  
QY 819 CAAGGACTTTATTGTCACCTTTGAGAAGGATCCGACGAGCGGTACACCTGTGAGAA 878  
Db 995 CAAGGACTTTATTGTCACCTTTGAGAAGGATCCGACGAGCGGTACACCTGTGAGAA 1054  
QY 879 GGCCTTGAGTCATCCCTGGATGAGGAAACACGGCCCTCCACGGGACATCTACCCATC 938  
Db 1055 GGCCTTGAGTCATCCCTGGATGAGGAAACACGGCCCTCCACGGGACATCTACCCATC 1114  
QY 939 AGTCAGCCTCCAGATCCAGAAAGCTTTGCTAAGAGCAAGTGGAGGCAAGCCTTCAAGC 998  
Db 1115 AGTCAGCCTCCAGATCCAGAAAGCTTTGCTAAGAGCAAGTGGAGGCAAGCCTTCAAGC 1174  
QY 999 AGCAGCTGGTGCACCATGAGGAGCTACACATGAACTGACACGCCCGGGGCTCCG 1058  
Db 1175 AGCAGCTGGTGCACCATGAGGAGCTACACATGAACTGACACGCCCGGGGCTCCG 1234  
QY 1059 CCCAGAGTGGAGAACAGAGCCGCTGAAACTCAAGCCTCAGAAAGCTTAGACCCAGCTC 1118  
Db 1235 CCCAGAGTGGAGAACAGAGCCGCTGAAACTCAAGCCTCAGAAAGCTTAGACCCAGCTC 1294  
QY 1119 CCCTGAGATACCATCAGCGAGGACCTGTCTCTGGACACAGTGTAGACATCTCCCTGGCCT 1178  
Db 1295 CCCTGAGATCAGCATCAGCGAGGACCTGTCTCTGGACACAGTGTAGACATCTCCCTGGCCT 1354  
QY 1179 GACCCAAATTACCTGCCAGCATGGCCGGCCCACTGCCCTGGTGGCAGGCTCCCTCAA 1238  
Db 1355 GACCCAAATTACCTGCCAGCATGGCCGGCCCACTGCCCTGGTGGCAGGCTCCCTCAA 1414  
QY 1239 CTGCCCTGGTCAATGGCTCCCTCCACATCAGCAGCAGCTGGTGGCCCATGATCAGGGGTC 1298  
Db 1415 CTGCCCTGGTCAATGGCTCCCTCCACATCAGCAGCAGCTGGTGGCCCATGATCAGGGGTC 1474  
QY 1299 CCTGGCCCGGGCCCTGTGGCTGTGCTCCAGCTGCCTGACATTTGGGAGCAAGGAA 1358  
Db 1475 CCTGGCCCGGGCCCTGTGGCTGTGCTCCAGCTGCCTGACATTTGGGAGCAAGGAA 1534  
QY 1359 GTCTCTACTGCTGAGCCACACTCCTCAAAAAGGCCAACAAAAACAGAACTTCAA 1418  
|||||

RESULT 5  
HS272L161

Db 1535 GTCTCTCTACTGCTGTGAGCCACACTCCTCAAAAAGGCCAACAAAAACAGAACTTCAA 1594  
QY 1419 GTCGGAGGTATGTTACCACTTAAAGCCOAGTGGCAGTCCCCTGCTCCGCGGAGGCGAC 1478  
Db 1595 GTCGGAGGTATGTTACCACTTAAAGCCOAGTGGCAGTCCCCTGCTCCGCGGAGGCGAC 1654  
QY 1479 TGGAGTCTGTCTCATTTATGTGATTTCTGGAGGCTGTGCCCTATGTCACTGCAATTTTCAGG 1538  
Db 1655 TGGAGTCTGTCTCATTTATGTGATTTCTGGAGGCTGTGCCCTATGTCACTGCAATTTTCAGG 1714  
QY 1539 AGACATATTCAACTCCTCTGCTTCCAAACCTGGTGTCTATCCGGCAGAGGGAAGG 1598  
Db 1715 AGACATATTCAACTCCTCTGCTTCCAAACCTGGTGTCTATCCGGCAGAGGGAAGG 1774  
QY 1599 CAGAGCAAGTGGAGCAGGCTTAGCAGGAGCAGTTTCTGGCCAGAGCACCACAGCTGTGTG 1658  
Db 1775 CAGAGCAAGTGGAGCAGGCTTAGCAGGAGCAGTTTCTGGCCAGAGCACCACAGCCGCTG 1834  
QY 1659 CAGCGGGGAGGCCCTCATAGGAGGCCAGAGGAGGCCCAAGCGTAGAAGCTTGT 1718  
Db 1835 CAGCGGGGAGGCCCTCATAGGAGGCCAGAGGAGGCCCAAGCGTAGAAGCTTGT 1894  
QY 1719 TGAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGGTCTCCCTGACCTGCCTGC 1778  
Db 1895 TGAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGGTCTCCCTGACCTGCCTGC 1954  
QY 1779 TCTATGCCCCACACCCTACGTGCCGTGTGTGCAAGTGTAGTACGTAGATGTCTCGGCT 1838  
Db 1955 TCTATGCCCCACACCCTACGTGCCGTGTGTGCAAGTGTAGTACGTAGATGTCTCGGCT 2014  
QY 1839 GGTCTGTGCTGTCTGCTGAAAGCTTAATGGGTGGCCAGGCTGTGTCACTTCTCC 1898  
Db 2015 GGTCTGTGCTGTCTGCTGAAAGCTTAATGGGTGGCCAGGCTGTGTCACTTCTCC 2074  
QY 1899 AAGCAAGCCATATGAGCATCTACCCAGACTCCCACTCTGCACACACTCACTCCCACT 1958  
Db 2075 AAGCAAGCCATATGAGCATCTACCCAGACTCCCACTCTGCACACACTCACTCCCACT 2134  
QY 1959 CTCAAGCTCCCACTCTTGGCAGATGGGCTCATTAATGTCTGCTGGCTGCCATCTGC 2018  
Db 2135 CTCAAGCTCCCACTCTTGGCAGATGGGCTCATTAATGTCTGCTGGCTGCCATCTGC 2194  
QY 2019 ATGAATGACAGGAGCTCCCACTGTGTGTCTGCTGTGAGCTCTTCAAGTCTTAATCTCT 2078  
Db 2195 ATGAATGACAGGAGCTCCCACTGTGTGTCTGCTGTGAGCTCTTCAAGTCTTAATCTCT 2254  
QY 2079 AACTCCAGATTAGCTCCCAAGTGGCTGAGACCCAGCAGCAGCAGCTTCTGCCCTTCTC 2138  
Db 2255 AACTCCAGATTAGCTCCCAAGTGGCTGAGACCCAGCAGCAGCAGCTTCTGCCCTTCTC 2314  
QY 2139 CTTGCTCTCAATCTAAAAGCAGTGCCACACCCCTCCAAAGTGGAAATAGAAAGAGTTTCA 2198  
Db 2315 CTTGCTCTCAATCTAAAAGCAGTGCCACACCCCTCCAAAGTGGAAATAGAAAGAGTTTCA 2374  
QY 2199 GTAAGGGCTGCAAGGAATTTCTATCTGCGCAGATGCTCCGTGGACACACCCCAATGGA 2258  
Db 2375 GTAAGGGCTGCAAGGAATTTCTATCTGCGCAGATGCTCCGTGGACACACCCCAATGGA 2434  
QY 2259 GTTAACTTGGAAAGTGTGACTTATTTTAAATGTCTGCCAGGAGTTCTAATCTGCTGTTC 2318  
Db 2435 GTTAACTTGGAAAGTGTGACTTATTTTAAATGTCTGCCAGGAGTTCTAATCTGCTGTTC 2494  
QY 2319 CTTTTTCTCTCTGAAAGTCCAGCACCAATTTCTGTCTCTTCCCCAGTTTCTCTCGCCT 2378  
Db 2495 CTTTTTCTCTCTGAAAGTCCAGCACCAATTTCTGTCTCTTCCCCAGTTTCTCTCGCCT 2554  
QY 2379 CCACCCCTCCAGCTTCTATGCTAGTGTGTGTCTAATAAATGGACATATTTTCTCT 2436  
Db 2555 CCACCCCTCCAGCTTCTATGCTAGTGTGTGTCTAATAAATGGACATATTTTCTCT 2612

LOCUS HS272L161 1738 bp mRNA linear PRI 21-APR-1999  
DEFINITION Human gene isolated from PAC 272L16, chromosome 1, similar to  
calcium/calmodulin dependent protein kinases.  
ACCESSION AL049688  
VERSION AL049688.1 GI:4678721  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1738)  
Rhodes,S.  
Direct Submission  
Submitted (21-Apr-1999) E-mail contact: humquery@sanger.ac.uk  
This sequence was generated from cDNA clones isolated using  
sequence from the bacterial clone 272L16 (AL023754) and EST data.  
The EST sequences listed match this sequence with an identity of at  
least 95% between the coordinates shown.  
Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr1/ Partial, experimentally  
determined gene.  
Sanger Centre name: dj272L16.Cl.1.

## FEATURES

Location/Qualifiers

1..1738  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="lg32.1-32.3"  
<1..1448  
/codon\_start=3

## CDS

/product="hypothetical protein"  
/protein\_id="CAB41259.1"  
/db\_xref="GI:467872"  
/translation="STLEAMGRKEEDDCSSWKQTTNIRKTFIMEVLGSGARSEVL  
VKQRTGLFALFKCIPKSPAPRDSLENIATVLAKIKHENITVLEDIYESTHYILVM  
QLVSGGELFLRILRGVYTKDASLVIOQVLSAVKYLHENGIVHRDLKPNLLYLP  
ENSKIMTDFGLSMEQIMSTACGTPGYVAPEVLAQPKYSKAVDCWSIGVITYILL  
CGPPFYDEETSKLFEKIEGYEFESPFWDIDISEAKDFICHILLEKDPNERTCEKA  
LSHPWIDNTALHRDIYPSVSLQIKNFASKRQAFNAAVVHHMKRLHMHNSPGV  
RPEVNRPEPTQASSTRSPSEITTEAPVLDHVALPALQLPCQGRPRPAPGR  
SLNCLVNSLHISVPMHOGSLAAGPCGCCSCLNIGSKGSSYCSSEPTLLKANK  
KONFKSEVMVPVKAGSGSHCRAGOTGVCLIN"  
<1..107

## exon

/number=1

108..236

## exon

/number=2

237..311

## exon

/number=3

312..450

## exon

/number=4

451..574

## exon

/number=5

575..650

## exon

/number=6

651..763

## exon

/number=7

653..777

## misc\_feature

/note="matches EST A1215131 from clone IMAGE:1925595"

## exon

764..842

## exon

/number=8

843..930

## exon

/number=9

868..1159

## misc\_feature

/note="matches EST AA351937"

## exon

931..1355

## exon

/number=10

1356..1483

## misc\_feature

/number=11  
Join(1469..1550,1553..1723,1721..1738)

## exon

/note="matches EST R05661 from clone 29560"

## exon

1484..1736

## exon

/number=12

## BASE COUNT

432 a 491 c 458 g 357 t

## ORIGIN

Query Match 70.9%; Score 1734.8; DB 9; Length 1738;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1736; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 53 CTTCAACTCTGGAGCAATGGTGCAGAAAGGAAGAAGTACTGCAGTTCTTGGAGAAAC 112  
DB 1 CTTCAACTCTGGAGCAATGGTGCAGAAAGGAAGAAGTACTGCAGTTCTTGGAGAAAC 60

QY 113 AGACCACCAACATCCCGAAACCTTCATTTTATGAAAGTCTGGGATCAGGAGCTTTCT 172  
DB 61 AGACCACCAACATCCCGAAACCTTCATTTTATGAAAGTCTGGGATCAGGAGCTTTCT 120

QY 173 CAGAAGTTTCTTGGTGAAGCAAGACTGACTGGGAAGCTCTTTTGTCTCTGAAGTCATCA 232  
DB 121 CAGAAGTTTCTTGGTGAAGCAAGACTGACTGGGAAGCTCTTTTGTCTCTGAAGTCATCA 180

QY 233 AGAAGTCACCTGCTTCCGGGACAGCAGCTGAGGAATGAGATTGCTGTGTGTTGAAAAGA 292  
DB 181 AGAAGTCACCTGCTTCCGGGACAGCAGCTGAGGAATGAGATTGCTGTGTGTTGAAAAGA 240

QY 293 TCAGCATGAAACATTTGACCTTGAGGACATCTATGAGACACACCCACTACTACC 352  
DB 241 TCAGCATGAAACATTTGACCTTGAGGACATCTATGAGACACACCCACTACTACC 300

QY 353 TGGTCATGCAGCTTGTCTTCTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCT 412  
DB 301 TGGTCATGCAGCTTGTCTTCTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCT 360

QY 413 ACAGAGAAGGATGCCAGTCTGTGTATCCAGCAGGTCTTGTGCGGAGTGAATACCTAC 472  
DB 361 ACAGAGAAGGATGCCAGTCTGTGTATCCAGCAGGTCTTGTGCGGAGTGAATACCTAC 420

QY 473 ATGAGAATGCATCGTCCACAGAGACTTAAAGCCCGAAACCTTACTTTACCCCTG 532  
DB 421 ATGAGAATGCATCGTCCACAGAGACTTAAAGCCCGAAACCTTACTTTACCCCTG 480

QY 533 AAGAGAATCTAAGATCATGATCAGTCTTGTGTCTGTCACAGATGGAACAGATGGCA 592  
DB 481 AAGAGAATCTAAGATCATGATCAGTCTTGTGTCTGTCACAGATGGAACAGATGGCA 540

QY 593 TCATGTCCACTGCTGTGGGACCCCGAGCTACGTGGCTCCAGAAAGTGTCTGGCCAGAAAC 652  
DB 541 TCATGTCCACTGCTGTGGGACCCCGAGCTACGTGGCTCCAGAAAGTGTCTGGCCAGAAAC 600

QY 653 CCTACAGCAAGGCTGTGGATTGTGTGTCCATCGCGGCTCATCATCATATTTCTGTG 712  
DB 601 CCTACAGCAAGGCTGTGGATTGTGTGTCCATCGCGGCTCATCATCATATTTCTGTG 660

QY 713 GATACCCCGCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAAGATCAAGAGGGCT 772  
DB 661 GATACCCCGCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAAGATCAAGAGGGCT 720

QY 773 ACTATGAGTTGAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGGCAAGGACTTTATTT 832  
DB 721 ACTATGAGTTGAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGGCAAGGACTTTATTT 780

QY 833 GCCACTTGTGTGAAGGATCCGAACGAGCGGTACACCTGTGAGAGGCTTTGAGTCATC 892  
DB 781 GCCACTTGTGTGAAGGATCCGAACGAGCGGTACACCTGTGAGAGGCTTTGAGTCATC 840

QY 893 CCTGGATTGAGGAAACACGCCCTCCACCGGACATCTACCCATCAGTCAGGCTCCAGA 952  
DB 841 CCTGGATTGAGGAAACACGCCCTCCACCGGACATCTACCCATCAGTCAGGCTCCAGA 900

QY 953 TCCAGAGAAGTCTTGTCTAAGAGCAAGTGGAGGCAAGCTTCAACGACGACGCTGTGTGTC 1012  
DB 901 TCCAGAGAAGTCTTGTCTAAGAGCAAGTGGAGGCAAGCTTCAACGACGACGCTGTGTGTC 960

QY 1013 ACCACATGAGGAAGCTACACATGAACCTGCACAGCCCGGGCGTCCGCCAGAGGTGGAGA 1072  
DB 961 ACCACATGAGGAAGCTACACATGAACCTGCACAGCCCGGGCGTCCGCCAGAGGTGGAGA 1020

```
QY 1073 ACAGGCGCGCTGAACCTCAAGCCTCAGAACCTCTAGAACCCAGCTCCCTCGATGATCACCA 1132
Db 1021 ACAGGCGCGCTGAACCTCAAGCCTCAGAACCTCTAGAACCCAGCTCCCTCGATGATCACCA 1080
QY 1133 TCACCGAGGACCTGCTCCCTGGACACAGTGTAGCACTCCCTGCCCTGACCCCAATTACCTT 1192
Db 1081 TCACCGAGGACCTGCTCCCTGGACACAGTGTAGCACTCCCTGCCCTGACCCCAATTACCTT 1140
QY 1193 GCCAGATGCGCCGCGCCCACTGCGCCCTGGTGGCAGTCCCTCAACTGCTCTGGTCAATG 1252
Db 1141 GCCAGATGCGCCGCGCCCACTGCGCCCTGGTGGCAGTCCCTCAACTGCTCTGGTCAATG 1200
QY 1253 GCTCCCTCCACATCAGCAGCAGCTGCTGCCATCATCAGGGTCCCTGGCGCGCGGC 1312
Db 1201 GCTCCCTCCACATCAGCAGCAGCTGCTGCCATCATCAGGGTCCCTGGCGCGCGGC 1260
QY 1313 CTTGTGCTGCTCTCCAGCTGCTGCAACATTGGGAGCAAGGAAAGTCTCTACTGCT 1372
Db 1261 CTTGTGCTGCTCTCCAGCTGCTGCAACATTGGGAGCAAGGAAAGTCTCTACTGCT 1320
QY 1373 CTGAGCCCACTCTTCAAAAAGGCCAACAAAACAGAACTTCAAGTCGGAGGTCATGG 1432
Db 1321 CTGAGCCCACTCTTCAAAAAGGCCAACAAAACAGAACTTCAAGTCGGAGGTCATGG 1380
QY 1433 TACCACTTAAAGCAGTGCAGCTCCCTGCTGCAACATTGGGAGCGGAGCTGGAGTCTGCTCA 1492
Db 1381 TACCACTTAAAGCAGTGCAGCTCCCTGCTGCAACATTGGGAGCGGAGCTGGAGTCTGCTCA 1440
QY 1493 TTATGTGATTCTCTGGAGCCTGCTGCTATGCTCACTGCAATTTTTCAGGAGACATATCAACT 1552
Db 1441 TTATGTGATTCTCTGGAGCCTGCTGCTATGCTCACTGCAATTTTTCAGGAGACATATCAACT 1500
QY 1553 CTTCTGCTCTTCAAACTGCTGCTATCCGCGAGGAGGAGGAGCAAGTGGAG 1612
Db 1501 CTTCTGCTCTTCAAACTGCTGCTATCCGCGAGGAGGAGGAGGAGCAAGTGGAG 1560
QY 1613 CAGGCGTTAGCAGGAGCAGTTTCTGGCCAGAGACACCGCTGCTGCCAGGGGCGAGCC 1672
Db 1561 CAGGCGTTAGCAGGAGCAGTTTCTGGCCAGAGACACCGCTGCTGCCAGGGGCGAGCC 1620
QY 1673 CCTCATAGAGGCGCCAGGAGGAGGAGCCCAAGGCGTAGAAGCCTTGTGAAGCTGTGAGCA 1732
Db 1621 CCTCATAGAGGCGCCAGGAGGAGGAGGAGCCCAAGGCGTAGAAGCCTTGTGAAGCTGTGAGCA 1680
QY 1733 GGAGAAGCGGTGCCACAGCTTCCAGGTCCTCCCTGACCTGCTCTATGCCCCAC 1790
Db 1681 GGAGAAGCGGTGCCACAGCTTCCAGGTCCTCCCTGACCTGCTCTATGCCCCAC 1738

RESULT 6
BC021840 2416 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, clone MGC:30513 IMAGE:4502479, mRNA, complete cds.
DEFINITION BC021840
ACCESSION BC021840.1 GI:18256866
VERSION MGC.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Strausberg, R.
Direct Submission
Submitted (18-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,  
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 41 Row: 1 Column: 13  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Genomescan gene  
prediction.

```
FEATURES
source
1..2416
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="MGC:30513 IMAGE:4502479"
/tissue_type="Eye, retina, mouse strain C57Bl/6"
/clone_lib="NIH_MGC_94"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
87..1520
/codon_start=1
/product="Unknown (protein for MGC:30513)"
/protein_id="AAH21840.1"
/db_xref="GI:18256867"
/translation="MGRKEEDCSHKKQTNIKTFFMEVLGSCAPEFLVKQV
TGKLPALCKIKKSPAFRDSLENIHAKRIKHENIVTLEDIEDSTHYLYVMQLVSG
GEILFDRLERGVYTEKDALVQQVLSAVKYLHENGIVHRDLKPENLLYLTPEENSKI
MITDFGLSKMEQNGVSTAGCTPGYVAPVLAQKPYSCADKWCSTGVITYILLGYPPI
FYBETESKLFKEIKEGYFSPFDDI2ESAKDFYCHLLEKDPNRYVCEKALRHPW
IDGNTALHRDLYPVSLOIOKNAFKSWROAFNAAVVHMKLHMHNLHSPKSRQVEE
NRPPVSPAPEVSRPDSITEAPILDPSTPLPALTRLPCSHSRPSAPSGRSLNC
LVNGSLRISLSLVPMOGGLATGPGCCSCSLNIGNKGSVCSEPTLFRKANKQNF
KSEVWPVVRAGGSHCRGGTGVCVLM"
BASE COUNT 572 a 694 c 602 g 548 t
ORIGIN
```

```
Query Match 54.3%; Score 1328.6; DB 10; Length 2416;
Best Local Similarity 76.1%; Pred. No. 0;
Matches 1853; Conservative 0; Mismatches 494; Indels 88; Gaps 14;
QY 31 GAGTCCCTGGCATCCTCAGAAAGCTTCAACTCTGGAGCAATGGTCCGAAAGGAAGAAT 90
Db 48 GAGCCCTGGCTTCTCAGGAGCTTCAACTCTGGAGCATGGGGCTAAGGAGGAGGAG 107
QY 91 GACTGCAGTTCCTGGAAGAACAGACACCACCAACATCCGGAACCTTCATTTTATGGAA 150
Db 108 GACTGCAGTTCCTGGAAGAACAGACACCACCAACATCAGGAACCTTCATCTTCATGGAA 167
QY 151 GTGCTGGGATCAGAGCTTCTCAGAAAGTTCCTGCTGTAAGCAAAAGACTGACTGGGAAG 210
Db 168 GTGCTGGGATCAGAGCTTCTCAGAGGTGTTCTGCTGTAAGCAAAAGACTGACTGGGAAG 227
QY 211 CTCTTTCTCTGAAGTGCATCAAGAGTCACTACCTGGTTCATGAGCTGTTCTGGTGGAGCATCTAT 270
Db 228 CTCTTTCTCTGAAGTGCATCAAGAGTCACTACCTGGTTCATGAGCTGTTCTGGTGGAGCATCTAT 287
QY 271 GAGATTCCTGTTGAAAGATCAAGCATGAAACATTTGACCCCTGGAGGACATCTAT 330
Db 288 GAGATTCCTGTTGAAAGATCAAGCATGAAACATTTGACCCCTGGAGGACATCTAT 347
QY 331 GAGAGCACCCCACTACTACTCTGGTTCATGAGCTGTTCTGGTGGAGGACATCTAT 390
Db 348 GAGAGCACCCCACTACTACTCTGGTTCATGAGCTGTTCTGGTGGAGGACATCTAT 407
QY 391 CGATTCCTGGAGGGGTGTCTACACAGAAAGATGCCAGTCTGGTGTATGCCAGAGGTC 450
Db 408 CGATTCCTAGAGCGTGTGTCTACACAGAAAGATGCCAGTCTGGTGTATGCCAGAGGTC 467
```



SOURCE Mus musculus.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 2427)  
AUTHORS Bjork, B.C., Watanabe, Y., Murray, J.C. and Schutte, B.C.  
TITLE Characterization of the human ortholog of rat Cam Kinase I gamma (CamK1g) at lq32-q41  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2427)  
AUTHORS Bjork, B.C., Watanabe, Y., Murray, J.C. and Schutte, B.C.  
TITLE Direct Submission  
JOURNAL Submitted (05-OCT-2001) Genetics, University of Iowa, 140 EMBR, Iowa City, IA 52242, USA  
FEATURES  
source Location/Qualifiers  
1..2427  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="1"  
/map="1H, 104.00 cM"  
/clone="IMAGE:4502479"  
/tissue\_type="retina"  
/note="corresponds to EST BG293660"  
gene 1..2427  
/gene="Camk1g"  
CDS 86..1519  
/gene="Camk1g"  
/codon\_start=1  
/product="calcium/calmodulin-dependent protein kinase I gamma"  
/protein\_id="AAL28101.1"  
/db\_xref="GI:16755794"  
translation="MGRKEEDCSWKKQTNIKRTIFMEVLGSGAFSEVFLVKQRY  
TGKLFALCKIKSPAFROSSLEIAVLKRIKENIVLEDIYESTHYLVMPOLVSG  
GELFOILLERGVYTEKSLVQVLSVAVLYHENGIVHRDLKPENLILYTPENSKI  
MTDFGLSKMEONGVNSTCGTGVVAPEVLAQRPYSKAVDCWSIGVITYILICGYPP  
FYEETSEKLFKEIKGYYEFESPFWDIDISEAKDFICHLLKDPNERYTCEKALRHPW  
IDGNTALHREYIPVSLQIOKNFAKSWROAFNAAVVHHMRKLHMLNLSPSVROEVE  
NRPVPSAPVRSRDPDSSIIEAPILDPPLPALTLPCSHSRSPSAPSGRSLNC  
LYNGSLRISSSLVPMOOGPLATGPGCGSSCLNIGNKSKSYCSEPIFLFRKANKQNF  
KSEVWPVPAKGGSTHCRGGGTGVLVM"  
BASE COUNT 584 a 694 c 601 g 547 t 1 others  
ORIGIN  
Query Match 54.3%; Score 1328.2; DB 10; Length 2427;  
Best Local Similarity 76.1%; Pred. No. 0;  
Matches 1852; Conservative 1; Mismatches 494; Indels 88; Gaps 14;  
QY 31 GAGTCCCTGGATCCTCAGAGCTTCAACTCTGGAGGCAATGGGTGCGAAAGGAAGAAT 90  
DB 47 GAGCCCTGGCTTCTCAGGAGCTTCAACTCTGGAGGACATGGGCGTAAAGGAGGAG 106  
QY 91 GACTGCAGTTCCTCGGAAGAACAGACACCAACATCCGGAACCTTCATTTTATGAA 150  
DB 107 GACTGCAGTTCCTCGGAAGAACAGACACCAACATCAGGAACCTTCATCTCATGGA 166  
QY 151 GTGCTGGATCAGGAGCTTCTCAGAGTTCCTGGTGAAGCAAGAGTCTGACTGGGAAG 210  
DB 167 GTGCTGGATCAGGAGCTTCTCAGAGTTCCTGGTGAAGCAAGAGTCTGACTGGGAA 226  
QY 211 CTCTTTGCTCAGTTCAGTTCAGAGTTCACCTGCTCCGGAGCAGCAGCTGGAGAA 270  
DB 227 CTCTTTGCTCAGTTCAGTTCAGAGTTCACCTGCTCCGGAGCAGCAGCTGGAGAA 286  
QY 271 GAGATTGCTGTGTTGAAAAGATCAAGCATGAAACATTTGACCTCGAGGACATCTAT 330  
DB 287 GAGATCGCTGTGTTGAAAAGATCAAGCATGAGACATTTGACCTCGAGGACATCTAT 346  
QY 331 GAGAGCACCAACCACTACTACCTGGTTCAGAGCTTGTTCGTGGGAGAGCTCTTTGAC 390  
DB 347 GAGAGCACCAACCACTACTACCTGGTTCAGAGCTTGTTCGTGGGAGAGCTCTTTGAC 406  
QY 391 CGGATCCTGGAGGGGGTGTCTACACAGAGAGGATGCCAGTCTGGTATCCAGCAGTCT 450

DB 407 CGGATCCTAGAGCGTGGTGTCTACACAGAAAGGATGCCAGCCTGTGTATCCAGCAGGTC 466  
QY 451 TTGTGCGCAGTGAATACCTACATGAGATGAGATGCGTCCACAGAGACTTAAAGCCCGAA 510  
DB 467 TTGTCTCGGTGAATACCTTCAATGAGATGCGTCCACAGAGATCTAAGCCTGAA 526  
QY 511 AACTGCTTTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCATGACTTTGTGTG 570  
DB 527 AACTGCTGTACCTACCCCTGAGGAGAACTCCAAGATCATGATCATGACTTTGTGTG 586  
QY 571 TCCAAGATGAACAGAGATGCATGCTCCACATGCTGTGGGACCCAGCTACGTCGCT 630  
DB 587 TCCAAGATGAGAGAGATGAGTCTATGTCACAGCTGTGGGACCCAGCTACGTCGCT 646  
QY 631 CCAGAGTGTGCGCCCAAGAAACCTTACAGAGGCTGTGATTTGTGTCCTACGTCGCT 690  
DB 647 CCAGAGTGTGCGCCCAAGAAACCTTACAGTAAAGCTGTGACTGTGCTGCTTGTGTC 706  
QY 691 ATCACTTACATATTGCTCTGTGGATACCCCGCTTCTATGAAGAAACGAGTCTAAGCTT 750  
DB 707 ATCACTTACATATTGCTGTGCTGTGCTATCCCGCTTCTATGAAGAAACGAGTCTT 766  
QY 751 TTGAGAGATCAAGGAGGCTTACTATGAGTGTGAGTCTCCATCTCTGGATGACATTTCT 810  
DB 767 TTTGAGAGATCAAGGAGGCTTACTAGAGTGTGAGTCTCCGCTTCTGGGATGACATTTCT 826  
QY 811 GAGTACGCCAAGGACTTTATTTGCCACTTGTCTTGAAGAGATCCGAAACGAGGCTACAC 870  
DB 827 GAGTACGCCAAGGATTTTATTTGCCATCTGCTGGAGAAAGGACCCAAACGAGGTACAC 886  
QY 871 TGTGAGAGGCTTGTGCTATCCTCGATGACGGAACACGCGCTCCACCGGGACATC 930  
DB 887 TGCAGAGAAACCTTCAGACACCCCTGGATTTGATGGAACACAGCCCTCCACCGGACATC 946  
QY 931 TACCCTTACGTCAGCTTCCAGATCCAGAGAACTTTGCTAAAGAGAAAGTGGAGGAAGCC 990  
DB 947 TATCATCTCTGTCAGCTTCCAGATTCAGAGAACTTTGCCAAGAGAAAGTGGAGGAAGCC 1006  
QY 991 TTCAGCCAGCAGCTGTGCTGACCATGAGGAAAGCTACACATGAACCTGCACAGCCG 1050  
DB 1007 TTCAGCCAGCTGCTGCTGCTATCATGAGGAAACTGCATGAACCTGCACAGCCG 1066  
QY 1051 GGGTCCGCGCAGAGGTGGAGAACAGCGCGCTGAAACTCAAGCCCTCAGAAACCTCTAGA 1110  
DB 1067 AGTGTCCGCAAGAGTGGAGAACAGGCGACCTGTGCTCCCTGCTCTGAGTCTCCAGA 1126  
QY 1111 CCCAGCTCCCTGAGATTCACATCCGAGGACCTGTCTCTGGACACACAGTGTAGCAGTC 1170  
DB 1127 CCAGACTCCCATGACAGCTCCATCAGAGGCGCCCATCTCTGGACCAAGACACACCCCTT 1186  
QY 1171 CCTGCTGACCCCAATTTACCTTCCAGCATGCGCGCGCGCCCTACGCCCTCCCTG 1227  
DB 1187 CTTGACATGACCCGACCTACCTGCTCAGACAGCTCCCGGCTTCAGCTCCAGTGGTGGC 1246  
QY 1228 AGTCCCTCAACTGCTGTGTCATATGCTCCCTCCACATCAGACAGCCTGCTGGTCCCATG 1287  
DB 1247 CGCTCACTCAACTGCTGTGTCACGCTCCCTGCGCATCAGCAGTAGTCTAGTGGCCATG 1306  
QY 1288 CATCAGGGGTCCCTGCGCGCGCGCGCTGTGGCTGTGCTGCTCCAGCTGCTTGAACATGGG 1347  
DB 1307 CAACAGGGGCCCCCTAGCCAGCGGGCCCCCTGTGGCTGTGCTGCTCCAGCTGTCTAATATGGG 1366  
QY 1348 AGCAAGGAAAGTCCCTCTACTGCTGTAGCCCAACACTCTCTCAAAAGGCCAACAAAAA 1407  
DB 1367 AACAAAGGAAAGTCTTCTCTACTGCTGTAGCCCTACCTCTCTCAGAAAGGCCAATAA 1426  
QY 1408 CAGAACTTCAAGTCGGAGGTCATGTTACAGTTAAAGCCAGTGGGAGCTCCACAGTCCG 1467  
DB 1427 CAGAACTTCAAGTCAGAGGTCATGTTACAGTCCAGTGGTGGAGCAGCAGCTCCG 1486  
QY 1468 GCAGGGCAGAGTCGAGTCTCTCATTTATGATTTCTGAGGCTGTGCTATGTCACATG 1527





Db 147711 CAGCTTCCAGGTCTCCCTGACCTGCGTCTATGCCCCACACCCCTAGCTGCGCGTGCTC 147770  
QY 1810 TGTGCAAGTGTACGTAGATAGCTCTGCCCTGGCTGTGCTGTGTTCTGCGTGAAGACTTAA 1869  
Db 147771 TGTGCAAGTGTACGTAGATAGCTCTGCCCTGGGTGTGCTGTGTTCTGCGTGAAGACTTAA 147830  
QY 1870 TGGGCTGGCCAGGCTGTGCACCTTCTCCAAGCAAGCCATATGGAGCATCTACCCAGAC 1929  
Db 147831 TGGGCTGGCCAGGCTGTGCACCTTCTCCAAGCAAGCCATATGGAGCATCTACCCAGAC 147890  
QY 1930 TCCCACTGTGCACACACTCACTCCCACTCTCAAGCTCCCAAGCTTGTGGCCAGATTTGGG 1989  
Db 147891 TCCCACTGTGCACACACTCACTCCCACTCTCAAGCTCCCAAGCTTGTGGCCAGATTTGGG 147950  
QY 1990 CTCAATTAAGTCTGCTGCGCCCATCTGCATGAATGACAGGAGCTGCCCATGGTGGTCT 2049  
Db 147951 CTCAATTAAGTCTGCTGCGCCCATCTGCATGAATGACAGGAGCTGCCCATGGTGGTCT 148010  
QY 2050 GCCTGTGAGCTTCTCAAGTCTTAATCCTTAACTTCCAGGATTAGCTCCCAAGTGCCTGAG 2109  
Db 148011 GCCTGTGAGCTTCTCAAGTCTTAATCCTTAACTTCCAGGATTAGCTCCCAAGTGCCTGAG 148070  
QY 2110 ACCAGCCAGCACACTCTGTGGCCCTTCTCCCTGCGCTCAATCTAAAGCAGTGCACACCC 2169  
Db 148071 ACCAGCCAGCACACTCTGTGGCCCTTCTCCCTGCGCTCAATCTAAAGCAGTGCACACCC 148130  
QY 2170 TCCAAAGTGAATAGAAAGAGTTCAATGAGTGAAGGCTCAAGGAATCTTATCTCGGCC 2229  
Db 148131 TCCAAAGTGAATAGAAAGAGTTCAATGAGTGAAGGCTCAAGGAATCTTATCTCGGCC 148190  
QY 2230 ACATGTCTCCCTCGTGCACACACCACTGAGTTAACTTGGAAAGTTGACATTTTAAATGTC 2289  
Db 148191 ACATGTCTCCCTCGTGCACACACCACTGAGTTAACTTGGAAAGTTGACATTTTAAATGTC 148250  
QY 2290 TGGCAGGAGTTCTAACTGCGCTCTGTTCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 2349  
Db 148251 TGGCAGGAGTTCTAACTGCGCTCTGTTCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 148310  
QY 2350 TTCTGTGCTTCCCTCGAGTTTCTCGCCCTCCACCCCTCCAGCTTTCATGCTCAGTGTGTG 2409  
Db 148311 TTCTGTGCTTCCCTCGAGTTTCTCGCCCTCCACCCCTCCAGCTTTCATGCTCAGTGTGTG 148370  
QY 2410 CTTAATAAATGACATATTTTCTCTA 2437  
Db 148371 CTTAATAAATGACATATTTTCTCTA 148398

RESULT 9  
D86557  
LOCUS D86557 Rattus norvegicus mRNA for Protein Kinase, partial cds. ROD 07-FEB-1999  
DEFINITION Rattus norvegicus mRNA for Protein Kinase, partial cds.  
ACCESSION D86557  
VERSION D86557.1 GI:2077933  
KEYWORDS Protein Kinase.  
SOURCE Rattus norvegicus embryo (E18) brain cDNA to mRNA, clone\_lib:S.  
Nakanishi clone:N5.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 1013)  
AUTHORS Yokokura,H., Terada,O., Naito,Y. and Hidaka,H.  
TITLE Isolation and comparison of rat cDNAs encoding  
JOURNAL Ca2+/calmodulin-dependent protein kinase I isoforms  
MEDLINE Biochim. Biophys. Acta 1338 (1), 8-12 (1997)  
REFERENCE 97228532  
AUTHORS 2 (bases 1 to 1013)  
Yokokura,H.  
TITLE Direct Submission  
JOURNAL Submitted (15-JUL-1996) Hisayuki Yokokura, Nagoya University School  
of Medicine, Department of Pharmacology; Tsurumai 65, Showa-ku,  
Nagoya, Aichi 466, Japan (Tel:052-744-2075, Fax:052-744-2083)  
FEATURES Location/Qualifiers

source  
1. .1013  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="N5"  
/tissue\_type="brain"  
/clone\_lib="S. Nakanishi"  
/dev\_stage="embryo (E18)"  
<1. .86  
87. .>1013  
/note="similar to calcium/calmodulin-dependent protein  
kinase I mRNA (L24907, L26288)"  
/codon\_start=1  
/product="Protein Kinase"  
/protein\_id="BAA19880.1"  
/db\_xref="GI:2077934"  
/translation="MGRKEEDCSWKQTNIRKTFIMEVLGSGAFSEFLVKQRV  
TGLFALKCIKSPAFRDSLENIAYLKRKHENIVLTLEDIYESTHYVLYMQLVSG  
GEULDRILERGVTYTKDASLVIOVLASVAKYHLENGIVHRDLKAPENLLLTPEENSKI  
MITDFGLSKMEQNGVSTACGTGYVAPEVLAQKPYSKAVDCWSIGVITILLCGTTP  
FYETESKLFKIKYGYEFSPWDDISAKDFICHLLLEKDPNERVYCEKALRHPW  
IDGNTALHRDIYPSVSIQIKNPAKSKWRQAFN"  
BASE COUNT 276 a 245 c 262 g 230 t  
ORIGIN  
Query Match 33.2%; Score 813; DB 10; Length 1013;  
Best Local Similarity 90.2%; Pred. No. 1.1e-213;  
Matches 870; Conservative 0; Mismatches 95; Indels 0; Gaps 0;  
QY 31 GAGTCCCTGGCATCCTCAGAGCTTCAACCTCGGAGCAATGGTCGAAAGGAGAGAT 90  
Db 48 GAGCCCTGGCTTCCCAGGAGCTTCAACCTCGGAGCATGGGGCGAAAGGAGGAG 107  
QY 91 GACTGCACTTCTCGAAGAAACAGACCCCAACATCCGGAAACCTTTCATTTATGAA 150  
Db 108 GACTGTAGCTCTGGAAGAACACACCCCAACATCAGGAACACCTTTCATCTCAT 167  
QY 151 GTGCTGGATCAGGAGCTTCTCAGAAGTTTTCCTGTGTAAGCAACACTGACTGGGA 210  
Db 168 GTGCTGGATCAGGAGGCTTCTCAGAGGTGTCTCGTGTGAAGCAAGAGTACTGG 227  
QY 211 CTCCTTGTCTGAAGTGCATCAAGAAGTCACTGCCCTCCGGGACACAGCTGGAGA 270  
Db 228 CTCCTTGTCTGAAGTGCATCAAGAAGTCACTGCCCTCCGGGACAGTACCTAGAG 287  
QY 271 GAGATTCGTGTGTTGAAAGATCAAGCATGAAACATTTGACCCCTGGAGGACATCT 330  
Db 288 GAGATTCGTGTGTTGAAAGATCAAGCATGAAACATTTGACCCCTGGAGGATATCT 347  
QY 331 GAGAGCACCCACTACTACTGCTCATGCAGCTTCTTCTGTTGGTGGAGCTCTTTGAC 390  
Db 348 GAGAGCACCCACTACTACTGCTCATGCAGCTTCTTCTGTTGGAGTGAGCTTTTTC 407  
QY 391 CGGATCCTGGAGCGGGTGTCTACACAGAGAAGATGCCAGTCTGGTGTATCCAGAGG 450  
Db 408 CGAATCCTAGAGCGTGGTGTCTACACAGAAAAGGATGCCAGTCTGTGTCATCCAG 467  
QY 451 TTGTCGGCAGTGAATACCTACATGAGAATGGCATGCTCCACAGAGACTTTAAAGCCG 510  
Db 468 TTGTCGGCAGTGAATACCTTTCATGAGAATGGCATGCTTTCACAGAGATCTAAAG 527  
QY 511 AACCTGCTTACCTTACCCCTGAAGAGAACCTTAAAGATCATGATCATGCTACTGCT 570  
Db 528 AACCTGCTTACCTTACCCCTGAGAGAACTCCAAGATATGATCATGCTACTGTTGG 587  
QY 571 TCCAAGATGGAACAGAAATGGCATCATGTCCACTGCCCTGTGGGACCCCGAGGCTAG 630  
Db 588 TCCAAGATGGAACAGAAATGGAGTTCATGTCCACAGCTTGTGGGACCCCGAGGCT 647  
QY 631 CCAGAGTGTGGCCCGCAAGAACCCCTACAGAGGCTGTGGATTCCTGCTCCATCGGCTC 690  
Db 648 CCAGAGTGTGGCCCGCAAGAACCCCTACAGTAAAGCTGTGGAGCTGTGCTCATTTG 707  
QY 691 ATCACCTACATATTGCTCTCTGTGGATACCCCGCTTCTATGAAGAAACGGAGTCTA 750



```
KEYWORDS      human.
SOURCE        Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 1671)
AUTHORS      Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B.,
              Abuin,A. and Sands,A.T.
TITLE        Novel human kinase proteins and polynucleotides encoding the same
JOURNAL      Patent: WO 0142433-A 5 14-JUN-2001;
              Lexicon Genetics Incorporated (US)
FEATURES     Location/Qualifiers
              source
                1..1671
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
BASE COUNT   453 a 437 c 448 g 333 t
ORIGIN
Query Match      20.6%; Score 504.6; DB 6; Length 1671;
Best Local Similarity 71.3%; Pred. No. 3.3e-128;
Matches 695; Conservative 0; Mismatches 274; Indels 6; Gaps 2;

Qy 66 GGCATGGTTCGAAGAGCAAGATGACGTCAGTTCCTGGAAGAACACAGACACCAACAT 125
Db 163 GGCCATGCCCGGAGAACCGGAGAGCAGCTCCTCTGGAAAGCAAGCTGAAGACAT 222

Qy 126 CCGGAAACCTTCAATTTTATGGAAGTGCTGGGATCAGGAGCTTCTCAGAAGTTTCTCT 185
Db 223 CAAGAAGTCTCGAGTTCAAGAGACCCTCGGAACCGGGCCCTTTTCGAAGTGTGTTT 282

Qy 186 GGTGAAGCAAGACTGACTGGGAGCTCTTGTCTGTAAGTGCATC---AAGAAGTCACC 242
Db 283 AGCTGAAGAGAAAGCAACTGGCAAGCTCTTGTCTGTAAGTGTATCCCTAAGAAAGCGCT 342

Qy 243 TGCCTTCGGGACAGCTGAGATGAGATTGCTGTTGAAAGATCAAGCATCA 302
Db 343 GAAGGCAAGAAAGCAGCATGAGATGAGATGAGCGCTCGGATGAGAAAGATTAAGCTGA 402

Qy 303 AAACATTGTGACCTGGAGACATCTATGAGAGCACACCCACTACTACTCTGGTCAATGA 362
Db 403 AAATATTGTTGCCCTGGAAGACATTTATGAAGCCCAAAATCACTGTACTTGGTCAATGA 462

Qy 363 GCTGTGTTCTGGTGGGAGCTTTGACCGGATCTTGGAGCGGGGTCTACACAGAGAA 422
Db 463 GCTGGTCTCGGTGGAGAGCTGTTGACCGGATAGTGGAGAGGGGTTTATACAGAGAA 522

Qy 423 GGATGCCAGTCTGGTCAGAGCTTGTGTCGGCAGTGAATACCTACATGAGAAATGG 482
Db 523 GGATGCCAGCACTCTGATCCGCCAAGTCTTGGACGCGGTACTACTATCTCCACAGAAATGG 582

Qy 483 CATGCTCCAGAGCTTAAGCCGGAACCTGCTTACCTTACCCCTGAGAGAACTC 542
Db 583 CATGCTCCAGAGACTCAAGCCCGGAAATCTTGTACTACAGTCAAGATGAGGAGTC 642

Qy 543 TAAGATCATGACTGACTTTGTCTGTCGAAGATGGAACAAATGG---CATCATGTC 599
Db 643 CAAATAATGATCAGTCTGATTTGATTTGATTAAGTGGAGGCAAGAGATGCTGATGTC 702

Qy 600 CACTGCTGTGGGACCCAGGCTACGTGGCTCCAGAAAGTGTCTGGCCCGCAAGACCCTACAG 659
Db 703 CACTGCTGTGGAACTCCAGGCTATGTGCTCTGAAAGTCTCTCGCCGCAAGAACTTACAG 762

Qy 660 CAAGGCTGTGATTTGCTGTCATCCGCTCATCACTACATATTTCTGTGGATPACCC 719
Db 763 CAAGCGCTGTACTGCTGTGTCATCGGAGTGATTTGCTTACTCTGCTCGGGCTATCCC 822

Qy 720 CCCGTTCTATCAAGAAACGGAGCTCTAAGCTTTTTCGAGAAGATCAAGAGGGCTACTATGA 779
Db 823 TCCCTTTTATGATGAAATGACTCCAAGCTCTTTTGACAGATCCTCAAGCGGGAATATGA 882

Qy 780 GTTTGAGCTCTCCATTCGCGGATGACATTTCTGAGTCAGCCAGGAGCTTTATTGGCACTT 839
Db 883 GTTTGACTCTCCCTACTCGGGATGACATCTCCGACTCTGCAAAAGACTTCATTTCGGGAACCT 942
```

```
Qy 840 GCTTCAGAGAGATCCGRACGAGCGGTACACTGTGAGAAGCGCTTGAGTCATCCCTCGAT 899
Db 943 GATGGAGAAGGACCCGAATAAAAGATACACGTGTGAGCAGCGACTCGGCACCCATCGAT 1002

Qy 900 TGAGGAAACACGCGCCTCCACCGGACATCTACCCATCAGTCAGCTCAGCCCTCAGATCCAGAA 959
Db 1003 CGCTGGTGACACAGCCCTCAACAAAACATCCACGAGTCCGTACAGCCCGCAGATCCGAA 1062

Qy 960 GAATTTGCTTAAGAGCAAGTGGAGGCAAGCCCTTCAAGCGAGCAGCTGTGGTGACACAT 1019
Db 1063 AAATTTTGCAGAGCAAAATGGAGCAAGCATTTAATGCCACGCGCGTCTGTGAGACATAT 1122

Qy 1020 GAGGAAGCTACACAT 1034
Db 1123 GAGAAAATACACCT 1137

RESULT 12
AX166517 AX166517 Sequence 8 from Patent WO0138503.
DEFINITION AX166517
ACCESSION AX166517
VERSION AX166517.1 GI:14546862
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 1074)
JOURNAL Flanagan,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,
        Novel human protein kinases and protein kinase-like enzymes
        Patent: WO 0138503-A 8 31-MAY-2001;
        Sugen, Inc. (US)
FEATURES Location/Qualifiers
          source
            1..1074
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 313 a 258 c 276 g 227 t
ORIGIN
Query Match      20.5%; Score 502.2; DB 6; Length 1074;
Best Local Similarity 71.3%; Pred. No. 1.4e-127;
Matches 692; Conservative 0; Mismatches 273; Indels 6; Gaps 2;

Qy 70 ATGGGTGCGAAAGAGAAAGATGACTGCGAGTTCCTGGAGAAACACAGACCAACATCCGG 129
Db 1 ATGGCCCGGAGAACGGCGAGAGCAGCTCCTCTGGAAAAAGCAAGCTGAAGACATCAAG 60

Qy 130 AAAACCTTCAATTTTATGGAAGTCTGGGATCAGAGCTTCTCAGAAGTTTCTCTGGTG 189
Db 61 AAGATCTTCGAGTTCAAGAGACCCCTCGGAACCGGGCCCTTTCCGAGTGGTTTASCT 120

Qy 190 AAGCAAAAGACTGACTGGGAAGCTTTGCTCTGAAGTGCATC---AAGAAGTCACTGCC 246
Db 121 GAAGAGAGGCAACTGGCAAGCTCTTGTCTGTAAGTGTATCCCTAAGAGCGCTCAAG 180

Qy 247 TTCGGGACAGCGCTGGAGAAATGAGATTGCTGTGTGAAAAAGATCAAGCATGAAAC 306
Db 181 GCAAGGAAAGACGATAGAGAATGAGATAGCCGTCTCTGAGAAAGATTAAGCATGAAAT 240

Qy 307 ATTGTGACCTTGGAGGACATCTATGAGAGCACCCACCTACTACTACCTGGTCACTGAGCTT 366
Db 241 ATTGTGCTTGGAGAGACATTTATGAAGCCCAATCACCTGTACTTGGTGTGACGCTG 300

Qy 367 GTTTCGTGGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGCTACACAGAGAAGAT 426
Db 301 GTGTCGGTGGAGAGCTGTTTGACCGGATAGTGGAGAAAGGGGTTTATACAGAGAGAT 360

Qy 427 GCCAGTCTGGTATCCAGCAGCTCTTCTGCGCAGTGAATACCTACATGAGATGGCATC 486
Db 361 GCCAGACTCTGATCCGCCAAGTCTTGGACGCGGTGTACTATCTCCACAGANTGGCATC 420
```

QY 487 GTCACAGAGACTTAAGCCGCAAAACCTGCTTTACCTTACCCCTCAAGAGAACTCTAAG 546  
DB 421 GTCCACAGAGACTCAAGCCGCAAAATCTCTGTACTACAGTCAAGATGAGGAGTCCAA 480  
QY 547 ATCATGATCACTGACCTTTGGTCTGTGCTCAAGATGGAACAGATGG---CATCATGTCAC 603  
DB 481 ATAATGATCAGTACCTTTGGATGCTCAAAATGGAGGGCAAAAGAGATGTGATGTCAC 540  
QY 604 GCCTGTGGGACCCAGGCTACGTGGCTCCAGAGTCTGTGCTGAGAGTCTGTGCTGATGATG 663  
DB 541 GCCTGTGGGACTCCAGCTATGCTGCTGAGTCTGTGCTGAGTCTGTGCTGAGTCTGTG 600  
QY 664 CTGTGTGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723  
DB 601 GCGCTTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
QY 724 TTTCTATGAAGAAACGGGAGTCTAAGCTTTTCGAGAGATCAAGAGGCTTACTATGAGTTT 783  
DB 661 TTTTATGATCAAAATGACTCCAAGCTTTTGGAGATCTCTCAAGGCGGAATATGAGTTT 720  
QY 784 GAGTCTCCATTTCTGGGATGACATTTCTGAGTCAAGCAAGGACTTTATTTGCTGCTGCT 843  
DB 721 GACTCTCCCTTACTGGGATGACATCTCCGACTCTGCAAAAGACTTCAATTCGGAACCTG 780  
QY 844 GAGAAGGATCCGACAGCGGTACACTCTGTGAGAGCGCTTGAGTCACTCCCTGGATTGAC 903  
DB 781 GAGAAGGACCCGAATAAAGATACAGTGTGAGCAGAGCTCGGCAACCATGATGCTGCT 840  
QY 904 GGAACACGCGCTCCACCGGGACATCTACCCATCAGTCAAGTCCAGATCCAGATCCAGAA 963  
DB 841 GGTGACACGCTCCACAAACATCCAGAGTCCGTCAGCGCCAGATCCGGAAGAAC 900  
QY 964 TTTGCTAAGAGCAAGTGGAGGCAAGCTTCAAGCAGCAGCTGTGCTGCAACCATGAGG 1023  
DB 901 TTTGCCAAGAGCAATGAGAGCAAGCAATTAATGTCACCGCGCTGCTGAGACATATGAGA 960  
QY 1024 AAGCTACACAT 1034  
DB 961 AACTACACCT 971

RESULT 13  
AX167587  
LOCUS AX167587 1074 bp DNA linear PAT 03-JUL-2001  
DEFINITION Sequence 3 from Patent WO0142435.  
ACCESSION AX167587  
VERSION AX167587.1 GI:14596987  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1074)  
AUTHORS Donoho, G., Scoville, J., Turner, C.A., Friedrich, G., Zambrowicz, B., Abuin, A., and Sands, A.T.  
TITLE Novel human kinase proteins and polynucleotides encoding the same  
JOURNAL Patent: WO 0142435-A 3 14-JUN-2001;  
LEXICON Lexicon Genetics Incorporated (US)  
FEATURES  
source  
location/Qualifiers  
1..1074  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 313 a 258 c 276 g 227 t  
ORIGIN  
Query Match 20.5%; Score 502.2; DB 6; Length 1074;  
Best Local Similarity 71.3%; Pred. No. 1.4e-127;  
Matches 692; Conservative 0; Mismatches 273; Indels 6; Gaps 2;  
QY 70 ATGGGTGCAAGGAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 129  
DB 1 ATGGCCGGGAGAGCGCGAGAGAGCTCTCTCTCTGGAAGAACGCAAGCTGAAGACATCAAG 60

QY 130 AAAACCTTCTATTTTGTGAAGTGTGGATCAGAGCTTTCTCAGAACTTTTCTCCTGGTG 189  
DB 61 AAGATCTTCAGTTCAAAGAGACCCCTCGGAACCGGGCCCTTTCCGAAGTGTGTTAGCT 120  
QY 190 AAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATC---AAGAAGTCACCTGCC 246  
DB 121 GAAGAGAGGCAACTGCAAGCTCTTTGCTGTGAAGTGTATCCCTAAGAAGCGCTGAAG 180  
QY 247 TTCGGGACAGCGCTGGAGAAATGAGATGCTGTGTTGAAAAAGATCAAGATCAAAAC 306  
DB 181 GCAAGAAAGACAGCATAGAGATGAGTAGCCGCTCTGAGAAAGATTAAAGCATGAAT 240  
QY 307 ATTGTGACCTGGAGGACATCTATGAGAGCACCACCCTACTACTGCTGCTGATGAGCTT 366  
DB 241 ATTGTTCCCTGGAAGACATTTATGAAGCCCAATCACTGTACTTGTGCTGATGAGCTG 300  
QY 367 GTTCTGCTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAGGAT 426  
DB 301 GTGTCGCTGGAGCTGTTTACCGGATGAGTAGAGAGAGGGTCTTATACAGAGAGGAT 360  
QY 427 GCGAGTCTGTGATPCCAGAGCTTGTGCGGAGTGAATACCTACATGAGAAATGGCATC 486  
DB 361 GCGAGCACTCTGATPCCAGAGCTTGTGGAGCGCTGTACTATCTCCACAGAAATGGCATC 420  
QY 487 GTCCACAGAGACTTAAGCCGCAAAACCTGCTTTACCTTACCCTCAAGAGACTCTAAG 546  
DB 421 GTCCACAGAGACTCAAGCCGCAAAATCTCTGTACTACAGTCAAGATGAGGATCCAA 480  
QY 547 ATCATGATCACTGACTTGTGCTGTCCAAGATGGAACAGAAATGG---CATCATGCTCACT 603  
DB 481 ATAATGATCAGTCACTTTGATGTTCAAAATGGAGGCAAGAGATGTGATGCTCACT 540  
QY 604 GCCTGTGGGACCCAGGCTACGTGGCTCCAGAGTGTGCGCCAGAAACCTTACAGCAAG 663  
DB 541 GCCTGTGGAACTCCAGGCTATGCTGCTCTGAAAGTCTCGCCAGAAACCTTACAGCAA 600  
QY 664 GCTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723  
DB 601 GCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
QY 724 TTTCTATGAAGAAACGGGAGTCTAAGCTTTTCGAGAAAGATCAAGAGGCTTACTATGAGTTT 783  
DB 661 TTTTATGATGAAATGACTCCAAGCTCTTTGAGCAGATCTCTCAAGGCGGAATATGAGTTT 720  
QY 784 GAGTCTCCATTTCTGGGATGACATTTCTGAGTCAAGCAAGGACTTTATTTGCCACTTGT 843  
DB 721 GACTCTCCCTTACTGGGATGACATCTCGGACTCTGCAAAAGACTTCAATTCGGAACCTG 780  
QY 844 GAGAAGGATCCGACAGCGGTACACTCTGTGAGAGCGCTTTGAGTCACTCCCTGGATTGAC 903  
DB 781 GAGAAGGACCCGAATAAAGATACAGTGTGAGCAGCAGCTCGGCAACCATGATGCTGCT 840  
QY 904 GGAACACGCGCTCCACCGGGACATCTACCCATCAGTCAAGTCCAGATCCAGATCCAGAA 963  
DB 841 GGTGACACGCTCCACAAACATCCAGAGTCCGTCAGCGCCAGATCCGGAAGAAC 900  
QY 964 TTTGCTAAGAGCAAGTGGAGGCAAGCTTCAAGCAGCAGCTGTGCTGCAACCATGAGG 1023  
DB 901 TTTGCCAAGAGCAATGAGAGCAAGCAATTAATGTCACCGCGCTGCTGAGACATATGAGA 960  
QY 1024 AAGCTACACAT 1034  
DB 961 AACTACACCT 971

RESULT 14  
AX167585  
LOCUS AX167585 1158 bp DNA linear PAT 03-JUL-2001  
DEFINITION Sequence 1 from Patent WO0142435.  
ACCESSION AX167585  
VERSION AX167585.1 GI:14596986  
KEYWORDS

SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1158)  
AUTHORS Donoho G., Scoville J., Turner C.A., Friedrich G., Zambrowicz B.,  
Abuin A. and Sands A.P.  
TITLE Novel human kinase proteins and polynucleotides encoding the same  
JOURNAL Patent: WO 0142435-A 1 14-JUN-2001;  
Lexicon Genetics Incorporated (US)  
FEATURES  
Location/Qualifiers  
1..1158  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 323 a 283 c 304 g 248 t  
ORIGIN  
Query Match 20.5%; Score 502.2; DB 6; Length 1158;  
Best Local Similarity 71.3%; Pred. No. 1.4e-127;  
Matches 692; Conservative 0; Mismatches 273; Indels 6; Gaps 2;  
Qy 70 ATGGGTGGAAGAGAGATGACTGCAGTTCCTGGAAGAAACAGACCACCAATCCGG 129  
Db 1 ATGGCCCGGAGAACGGGAGAGCAGCTCCTCTGGAAGAAAGAGCTGAAGACATCAAG 60  
Qy 130 AAACCTTCATTTTATGGAAGTCTGGGATCAGGACCTTTCAGAACTTTTCCTGGTG 189  
Db 61 AAGATCTTCGAGTTCAAGAGACCCCTCGGAACCGGGGCCCTTTCCGGAAGTGGTTTAGCT 120  
Qy 190 AAGCAAGACATGACTGGGAAGCTTTTGCTCTGAAGTGCATC---AAGAAGTCACCTGGCC 246  
Db 121 GAAGAGAAGCACTGGCAAGCTCTTGCTGTGAAGTGTATCCCTAAGAGGCGCTGAAG 180  
Qy 247 TTCGGGACACGCTGGAGAAATGAGATGCTGTGTGAAAAGATCAGCATGAAGAAC 306  
Db 181 GGCAAGGAACAGCAGATAGAAATGAGATAGCCGCTCTGAGAAAGATTAAGCATGAAGAT 240  
Qy 307 ATTGTGACCTTGGAGGACATCTATGAGAGCACCAACCTACTACCTGCTGTCAGCAGCTT 366  
Db 241 ATTGTGCTTGGAGACATTTATGAAGCCCAATCACCTGTACTTGGTTCATGCACTG 300  
Qy 367 GTTCTGTGGGGAGCTCTTTGACCGGATCTTGAGCGGGGTGTCTACAGAGAAGGAT 426  
Db 301 GTGTCCGGTGAGAGCTGTTTGACCGGATAGTGAGAGGGGGTTTTATACAGAGAAGGAT 360  
Qy 427 GCCAGTCTGTGATCCAGCAGCTCTGTGCGCAGTGAATACCTACATGAGATGGCATC 486  
Db 361 GCCAGACTGTATCCCGCAAGTCTTGACCGCGTGTACTATCTCCACAGAATGGGCATC 420  
Qy 487 GTCCACAGAGACTTAAAGCCGAAACCTGCTTTTACCTTACCCTGAAGAACTCTTAAG 546  
Db 421 GTCCACAGAGACCTCAAGCCGAAATCTCTTGTTACTACAGTCAAGATGAGGATCCAAA 480  
Qy 547 ATCATGATCATGACTTTGGTCTGTCCAAGATGGAACAGAAATGG---CATCATGTCCACT 603  
Db 481 ATAATGATCATGACTTTGGATTGTCAAAAATGAGGCGCAAGAGAGATGTGATGCCACT 540  
Qy 604 GCCTGTGGAGCCAGGCTAGTGGCTCCAGAGTGTGCGCCAGAAACCTACAGCAAG 663  
Db 541 GCCTGTGGAACCTCAGGCTATGTGCTCTTGAAGTCTCGCCCAAGAACTTACAGCAA 600  
Qy 664 GCTGTGATTCGTGTCATCCGCGTCAATCACCTACATATTTGCTCTGTGGATACCCCGG 723  
Db 601 GCGGTTGACTGCTGTCCATCGGAGTGATTGCCCTACATCTTGCTCTCGGCTACCCCTCT 660  
Qy 724 TTCTATGAAGAAACGGAGCTTAAGCTTTTCGAGAAGATCAAGAGGGCTACTATGAGTTT 783  
Db 661 TTTTATGATGAATGACTTCAAGCTCTTTGAGCAGATCTCTCAAGGCGGAATATGAGTTT 720  
Qy 784 GAGTCTCATTCTGGGATGACATTTCTGAGTCAGCCCAAGGACTTTATTTGCACTCTGCTT 843  
Db 721 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAGACTTCATTCGGAACCTGATG 780

Qy 844 GAGAAGATCCGAACGAGCGGTACACCTGTGAGAAGGCCCTTGAGTCATCCCTGGATTGAC 903  
Db 781 GAGAAGGCCCGCCCAATAAAGATACACGTGTGACGACAGCTCGGCACCCATGGATCGCT 840  
Qy 904 GGAACACGCGCCTCCACCGGACATCTACCCATCAGTCAGCTCCAGATCCAGAAGAAC 963  
Db 841 GTGACACAGCCCTCAACAAAACATCCACGAGTCCGTCAGGCCAGATCCGGAAGAAC 900  
Qy 964 TTGTGTAAGCAAGTGGAGCAAGCCTTCAACGACGACGCTGTGTGTGCACCATAGAG 1023  
Db 901 TTGTCACAGAGCAATGAGAGCAAGCATTTAATGCCACGCGCTGCTGAGACATATGAGA 960  
Qy 1024 AAGCTACACAT 1034  
Db 961 AACTACACCT 971  
RATCAMPKAA 1439 bp mRNA linear ROD 18-MAY-1995  
Rattus norvegicus Cam-like protein kinase mRNA, complete cds.  
L26288  
VERSION L26288.1 GI:439613  
KEYWORDS protein kinase.  
SOURCE Rattus norvegicus (strain Sprague-Dawley) (library: day 21  
gestation lung cDNA) fetal lung cDNA to mRNA.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 1439)  
AUTHORS Cho, F.S., Phillips, K.S., Bogucki, B. and Weaver, T.E.  
TITLE Characterization of a rat cDNA clone encoding  
JOURNAL calcium/calmodulin-dependent protein kinase I  
MEDLINE Biochim. Biophys. Acta 1224 (1), 156-160 (1994)  
PUBMED 95035115  
FEATURES 7948038  
Location/Qualifiers  
1..1439  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/tissue\_type="lung"  
/dev\_stage="fetal"  
/tissue\_lib="day 21 gestation lung cDNA"  
118..1242  
/codon\_start=1  
/product="Cam-like protein kinase"  
/protein\_id="AAA66944.1"  
/db\_xref="GI:439614"  
/translation="MPGAVEGPRNKAQEDIRIDYDFRDLVLTGAFSEVILAEKRTQK  
LVAIKCIAKALEGSGSMENFIAVLHKIHPNIVALDDIYESGHLVLIIMQLVSGGE  
LFDRIVEKGYTERDASRLIFQVLDVAVYLDLGIYHRLDKPENLLYYSLDEDSKIMI  
SDFGLSKMEDPGLSVLTACGTPGYVAPVLAQKPYSKAVDCWSGYVAYILLCGTPPF  
YDENDAKLIFEQILKAEYBFDSPDDISDAKDFIRHLMKDKPKRFTCEALQHPWI  
AGDTALDKNIHQSVSEQIKKNFAKWKQAFNATAVVRMRKQLQGTSGEQGGQTASH  
GBLLPTAGGPAAGCCCRDCCVCPSELPPPPSSRAM"  
BASE COUNT 357 a 375 c 405 g 302 t  
ORIGIN  
Query Match 20.1%; Score 493; DB 10; Length 1439;  
Best Local Similarity 71.7%; Pred. No. 5.2e-125;  
Matches 675; Conservative 0; Mismatches 260; Indels 6; Gaps 2;  
Qy 101 CCTGGAAGAAACAGACCACCACATCCGAAACCTTCATTTTATGGAAGTCTGGGAT 160  
Db 140 CCAGGTGGAAGCAGGAGAGACATTTAGGATATTTATGACTTCAGGAGTTCTGGGCA 199  
Qy 161 CAGGAGCTTCTCAGAAAGTTTCTCGTGGTGAAGCAAGAGACTGACTGGAAAGCTCTTTGCTC 220  
Db 200 CGSGTGCCTTCTCAGAAAGTGATCCTGGCAGAGGACAAAGAGGACTCAGAAACTGGTGCCA 259  
Qy 221 TGAAGTGCAT---CAAGAAGTCACTGCGCTTCGGGACAGCAGCCCTGGAGAATGAGATTG 277

Search completed: March 14, 2003, 17:25:23  
Job time : 7400 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 14, 2003, 17:36:01 ; Search time 328 Seconds  
(without alignments)  
3268.144 Million cell updates/sec

Title: US-09-960-643-2

Perfect score: 2513  
Sequence: 1 MGRKEDDCSSWKQTTNIR.....VKAGSGSHCRAGTGVCLIM 476

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09960643/runat\_07032003\_090517\_19867/app\_query.fasta\_1.647  
-DB=N\_Geneseq\_101002 -OFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USBR=US09960643.cgn\_1\_1\_263@runat\_07032003\_090517\_19867 -NCFU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	2513	100.0	1956	22	AAI60703	Human polynucleoti
2	2513	100.0	2447	24	AAD36140	Human calmodulin k
3	2499.5	99.5	2165	22	AAI58917	Human polynucleoti
4	2327.5	92.6	2689	22	AAS31014	Human diagnostic a
5	1250.5	49.8	1736	22	AAD18817	Human kinase (PRIN
6	1246	49.6	1074	22	AAS06708	Polynucleotide seq
7	1246	49.6	1074	22	AAH25119	Nucleotide sequenc
8	1246	49.6	1158	22	AAH25118	Nucleotide sequenc
9	1246	49.6	1565	24	ABL60905	Human CAM kinase I
10	1246	49.6	1661	22	AAH18826	Human kinase (PRIN
11	1246	49.6	1671	22	AAH25120	Nucleotide sequenc
12	1246	49.6	2164	22	AAI60424	Human polynucleoti
13	1243	49.5	1733	22	AAI58638	Human polynucleoti
14	1238.5	49.3	1694	22	AAC90432	Murine Dendritic C
15	1237	49.2	1578	22	AAAD04775	Human death domain
16	1150.5	45.8	1728	21	AAZ46162	cDNA sequence enco
17	1070	42.6	1458	19	AAV18867	Calmodulin-depende
18	1041	41.4	1032	22	AA06711	Polynucleotide seq
19	1036	41.2	1512	23	AA070157	DNA encoding novel
20	990	39.4	1282	20	AA06835	Disease associated
21	957.5	38.1	2264	23	ABL02719	Drosophila melanog
22	753	30.0	2494	22	AAK94614	Human full-length
23	753	30.0	2747	22	AAI58143	Human polynucleoti
24	753	30.0	2897	22	AAF30481	Human protein phos
25	753	30.0	2905	22	AAK94325	Human full-length
26	747.5	29.7	2297	22	AAH78261	Nucleotide sequenc
27	747.5	29.7	3017	22	AAH18832	Human kinase (PRIN
28	747.5	29.7	3124	24	ABA95682	Human protein kina
29	747	29.7	2840	22	AAI58144	Human polynucleoti
30	733.5	29.2	3022	22	AAK94299	Human full-length
31	726.5	28.9	1572	22	AAH99632	Human protein enco
32	683.5	27.2	1774	23	AA68767	DNA encoding novel
33	683.5	27.2	1774	23	AA090919	DNA encoding novel
34	667	26.5	1763	23	AA066250	DNA encoding novel
35	664.5	26.4	3471	18	AAAT71761	PSKH-1 cDNA. Homo
36	661.5	26.3	1764	23	AA064534	DNA encoding novel
37	650.5	25.9	1386	22	AAK52586	Human polynucleoti
38	650.5	25.9	1450	22	AAK51602	Human polynucleoti
39	639.5	25.4	3552	24	ABL40502	Human serine/threo
40	633.5	25.2	2867	24	AAD34311	Human PKIN-14 cDNA
41	625.5	24.9	2301	24	ABL40503	Human serine/threo
42	615.5	24.5	2110	21	AAZ46160	cDNA sequence enco
43	611.5	24.3	1803	22	AAH46903	cDNA encoding huma
44	611.5	24.3	1803	24	AAD38861	Human kinase (PRIN
45	611.5	24.3	7626	24	ABT04063	Human ovary specif

ALIGNMENTS

RESULT 1

AAI60703  
ID AAI60703 standard; cDNA; 1956 BP.  
XX AAI60703;  
AC AAI60703;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 4692.  
XX  
KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX  
OS Homo sapiens.

XX WO200153312-A1.  
XX 26-JUL-2001.  
XX 26-DEC-2000; 2000WO-US34263.  
XX 21-JAN-2000; 2000US-0488725.  
XX 25-APR-2000; 2000US-0552317.  
XX 09-JUL-2000; 2000US-0598042.  
XX 19-JUL-2000; 2000US-0620312.  
XX 03-AUG-2000; 2000US-0653450.  
XX 14-SEP-2000; 2000US-0662191.  
XX 19-OCT-2000; 2000US-0693036.  
XX 29-NOV-2000; 2000US-0727344.  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI: 2001-442253/47.  
XX P-PSDB; AAM41547.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
XX Claim 1; SEQ ID NO 4692; 10078pp; English.  
XX  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
XX Sequence 1956 BP; 474 A; 559 C; 508 G; 414 T; 1 other;  
SQ

Alignment Scores:  
Pred. No.: 5.24e-206 Length: 1956  
Score: 2513.00 Matches: 476  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 22 Gaps: 0

US-09-960-643-2 (1-476) x AAI60703 (1-1956)

QY 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20  
DB 65 ATGGGTGGAAGGAGCAAGATGACTGCTCTCTGGAAGAAACAGACCACCAATCCGG 124  
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
DB 125 AAAACCTTCAATTTTATGGAAGTGTCTGGGATCAGGAGCTTCTCAGAAAGTTTCTCGTGTG 184  
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysSerProAlaPhe 60  
DB 185 AAGCAAGAGTACTGCTGGGAGGCTCTTTGCTCTGAAAGTGCATCAAGAGTCACTGCCCTTC 244  
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80

Db 245 CGGCACAGCAGCCTGGAGAAATGAGATTGCTGTGTAAGAAAGATCAAGCATGAAAAACATT 304  
QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100  
Db 305 GTGACCTCTGGAGGACATCTATGAGAGCACCACTACTACTCTGCTGTCAGCTGCTGTT 364  
QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120  
Db 365 TCTGTTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGGATGCC 424  
QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140  
Db 425 AGTCTGTGTATCCAGCAGCTGTTGTGGCAGTGAATACCTACTAGATGAGATCGATCGTC 484  
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160  
Db 485 CACAGAGACTTAAGCCCGAAACCTGCTTTACCTTACCTCCCTGAAGAGAATCTAAGATC 544  
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180  
Db 545 ATGATCACTGACTTTGCTGTCTCCAGATGGAACAGAAATGGCATCATGTCCACTGCTGT 604  
QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200  
Db 605 GGGACCCAGGCTACGTGGCTCCAGAACTGCTGGCCAGAAACCTACACAGAGGCTGTG 664  
QY 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyr 220  
Db 665 GATTGCTGGTCCATCGGCGTCATCACCTACATATTGCTGTGGATACCCGCCATCTCTAT 724  
QY 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240  
Db 725 GAAGAAACGAGTCTAAGCTTTTCGAGAGATCAAGAGGGCTACTATGATGTTGAGTCT 784  
QY 241 ProPheTrpAspIleSerGluSerAlaLysAspPheIleCysHisLeuGluLys 260  
Db 785 CCATCTGGGATGACATTTCTGAGTCAGCAAGGACTTTATTTGCCACTTGTGTGAGAAG 844  
QY 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280  
Db 845 GATCCGAACGAGCGGTACAGCTGTGAGAGGCTTTGAGTATCCCTGGATTCAGCGAAAC 904  
QY 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300  
Db 905 ACACCCCTCCACCGGGACATCTACCCATCAGTCAGCTCCAGATCCAGAGAACTTTGCT 964  
QY 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu 320  
Db 965 AAGAGCAAGTGGAGGCAAGCCTTCAAGCAGCAGCTGTGTCACCAACATGAGGAAGCTA 1024  
QY 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProProGluThr 340  
Db 1025 CACATGAACCTGCAAGCGCGGCGTCCGCCAGAGGTGGAGAACAGCGCGCTGAAACT 1084  
QY 341 GlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaProVal 360  
Db 1085 CAACCTCAGAAACCTCTAGACCAGCTCCCTGAGATCACCATCAGGAGGACCTGTC 1144  
QY 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380  
Db 1145 CTGCACCACAGTGTAGCAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1204  
QY 381 ProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400  
Db 1205 CCCACTGCCCTGTGTGAGGTCCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1264  
QY 401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420  
Db 1265 AGCAGCTGTGTGCCCATGCATCAGGGGTCCCTGCGCCCGGGCGCTGTGCTGCTGCTGCT 1324  
QY 421 SerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCysSerGluProThrLeuLeu 440  
Db 1325 AGCTGCTGAACATTGGGAGCAAGAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1384

QY 441 LysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAlaSer 460  
Db 1385 AAAAAGCCCAACAAAACAACTTCAAGTCGGAGGTCATGTACCACTAAAGCCAGT 1444  
QY 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476  
Db 1445 GGCAGCTCCACCTGCCGGGAGGCAGACTGGAGTCTGTCTCATATG 1492

RESULT 2  
AAD36140  
ID AAD36140 standard; DNA; 2447 BP.  
AC AAD36140;  
XX  
XX 09-AUG-2002 (first entry)  
XX Human calmodulin kinase, CAMK-X1 gene.  
XX  
XX Human; cytostatic; antisense gene therapy; screening; protein kinase;  
XX cancer; liver; colon; tumour; inflammation; arthritic synovium; CAMK-X1;  
XX calmodulin kinase; enzyme; gene; chromosome lq32.1-32.3; ds.  
XX Homo sapiens.  
XX  
XX Location/Qualifiers  
XX Key 70..1500  
XX CDS /\*tag= a  
XX /product= "Human CAMK-X1 protein"  
XX  
XX WO200224947-A2.  
XX  
XX 28-MAR-2002.  
XX  
XX 20-SEP-2001; 2001WO-IB02237.  
XX  
XX 20-SEP-2000; 2000US-233999P.  
XX 02-OCT-2000; 2000US-237419P.  
XX 02-OCT-2000; 2000US-237423P.  
XX 04-OCT-2000; 2000US-238558P.  
XX 10-MAY-2001; 2001US-290555P.  
XX (KINE-) KINETEK PHARM INC.  
XX (UYBR-) UNIV BRITISH COLUMBIA.  
XX Yoganathan T, Delaney AD;  
XX  
XX WPI; 2002-394145/42.  
XX P-PSDB; AAE22764.  
XX  
XX Diagnosing cancer, comprises determining the upregulation of expression  
XX of a nucleic acid sequence encoding a protein kinase or upregulation of  
XX expression of the protein kinase, in the cancer -  
XX  
XX Claim 16; Page 62-64; 87pp; English.  
XX  
XX The invention relates to a method for screening biologically active agent  
XX that modulates cancer associated protein kinase function. The invention  
XX also relates to a method for diagnosing cancer comprising determining the  
XX upregulation of expression of a nucleic acid sequence encoding a protein  
XX kinase. The method is useful for diagnosing cancer. A protein kinase is  
XX useful for screening biological agents that modulate cancer associated  
XX protein kinase function. Downregulating the activity of protein kinase is  
XX useful for inhibiting the growth of a cancer cell, e.g. liver or colon  
XX cancer. A nucleic acid encoding protein kinase is useful to screen biopsy  
XX derived tumours and inflammatory samples such as arthritic synovium, for  
XX amplified DNA in the cell or increased expression of corresponding mRNA  
XX or protein and is also useful to detect differences in expression levels  
XX such as molecular weight, amino acid and nucleotide sequences between the  
XX two cells. The present sequence is human calmodulin kinase CAMK-X1 gene  
XX located on chromosome lq32.1-32.3.  
XX  
XX Sequence 2447 BP; 590 A; 707 C; 604 G; 546 T; 0 other;

Alignment Scores:  
Pred. No.: 7.2e-206 Length: 2447  
Score: 2513.00 Matches: 476  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0  
US-09-960-643-2 (1-476) x AAD36140 (1-2447)  
QY 1 MetGlyArgLysGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20  
Db 70 ATGGTCGAAAGGAAGAGATGACTGCAGTTCCTGGAGAAACAGACACCAATCCGG 129  
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
Db 130 AAAACCTTCATTTTATGGAAGTCTGGATCAGGAGCTTCTCAGAAGTTTCTCGGTG 189  
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60  
Db 130 AAGCAAGACATGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAAGTCACTGCCTTC 249  
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80  
Db 250 CGGACACAGCACCICGGAGATGAGATTGCTGTGTGAAAGATCAAGACATGAACATT 309  
QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100  
Db 310 GTGACCTGGAGGACATCTATGAGACGACACCCACCTACTACCTGGTGCATGCAGCTTGT 369  
QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120  
Db 370 TCTGGTGGGAGCTCTTTGACCCGATCTGGAGCGGGTGTCTACACAGAGAAAGATGCC 429  
QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140  
Db 430 AGTCTGGTGATCCAGCAGGTCTTCTCGCAGTGAATACCTACATGAGATGGCATCGTC 489  
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160  
Db 490 CACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCCTGAAGAGAACTCTAAGATC 549  
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180  
Db 550 ATGATCACTGACTTGGTCTGTCCAGATGGAACAGATGGCATCATGTCCACTGCCCTGT 609  
QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200  
Db 610 GGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCCCAGAAACCCCTACAGCAAGGCTGTG 669  
QY 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProProPheTyr 220  
Db 670 GATTGCTGGTCCATCGGCGTCATCACCTACATATGTCTGTGGATACCCCGCTTCAT 729  
QY 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240  
Db 730 GAACAAAGGAGCTAAGCTTTTCGACAAGATCAAGGAGGCTACTATGATTTGAGTCT 789  
QY 241 ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260  
Db 790 CCATCTCTGGGATGACATTTCTGAGTCAGCCCAAGGACTTTATTGGCCACTTGTGTGAAG 849  
QY 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280  
Db 850 ATCCGAAACGAGCGGTACCTGTGAGAAGCCCTTGTAGTTCATCCCTGGATGAGCGAAG 909  
QY 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300  
Db 910 ACGGCCCTCCACCGGACATCTACCCATCAGTCAGCCCTCCAGATCCAGAGAACTTTGCT 969  
QY 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu 320  
|||||

Db 970 AAGAGCAAGTGGAGGCAAGCCCTTCAACGACGAGCTGTGGTGCACACCATGAGGAAGCTA 1029

Qy 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProGluThr 340

Db 1030 CATATGAACCTGCACAGCCGGCGCTCCGCCAGAGGTGGAGAACAGGCCGCTGAACCT 1089

Qy 341 GlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaProVal 360

Db 1090 CAAGCCTCAGAAACCTCTAGACCAGCTCCCTGAGATCACCATCACCAGGACCTGCTC 1149

Qy 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380

Db 1150 CTGGACACAGGTAGACATCCCTGCGCTGACCCCAATTACCTGGCAGATGGCGCGGG 1209

Qy 381 ProThrAlaProGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400

Db 1210 CCCACTGCCCTGGTGGCAGGTCCCTCACTGCTGGTCAATGGTCCCTCCACATCAGC 1269

Qy 401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420

Db 1270 AGCAGCTGTGCCATGCATCAGGGGTCCCTGGCGCGCGGCGCTGTGGCTGCTGCTCC 1329

Qy 421 SerCysLeuAsnIleGlySerGlyLysSerSerTyrCysSerGluProThrLeuLeu 440

Db 1330 AGCTGCTGAACATGGGAGCAAGAAAGTCTCTACTGCTCTGAGCCACACTCCTC 1389

Qy 441 LysLysAlaAsnLysLysGlnAsnPhelLysSerGluValMetValProValLysAlaSer 460

Db 1390 AAAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTCTATGTACCATGTTAAAGCCAGT 1449

Qy 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476

Db 1450 GGCAGCTCCCACTGCCGGCAGGCGCAGACTGGAGTCTGTCTCATTTATG 1497

RESULT 3

ID AAI58917 standard; cDNA: 2165 BP.

XX AAI58917;

XX 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 1120.

DE Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

XX leukaemia; ss.

XX Homo sapiens.

OS W0200153312-A1.

PN 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

DR P-PSDB; AAM39761.

XX Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

XX Claim 1; SEQ ID NO 1120; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and

CC the encoded polypeptides (AAM38642-AAM42213) with neotropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

XX SQ Sequence 2165 BP; 533 A; 613 C; 563 G; 455 T; 1 other;

Alignment Scores:

Pred. No.: 8.74e-205 Length: 2165

Score: 2499.50 Matches: 476

Percent Similarity: 98.55% Conservative: 0

Best Local Similarity: 98.55% Mismatches: 0

Query Match: 99.46% Indels: 7

DB: 22 Gaps: 1

US-09-960-643-2 (1-476) x AAI58917 (1-2165)

Qy 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20

Db 253 ATGGTCCGAAAGGAAGATGACTGCAGTTCTTCTGGAAGAACAGACCACCATCCGG 312

Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40

Db 313 AAAACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTTCTCAGAAGTTTCTCTGGT 372

Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60

Db 373 AAGCAAGACCTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAGTCACCTGCCTTC 432

Qy 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80

Db 433 CGGGACAGCAGCTGGAGAATGAGATTGCTGTGTGAAAAAGATCAAGCATGAAACATT 492

Qy 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100

Db 493 GTGACCTGGAGGACATCTATGAGAGCACCCACTACTACCTGGTTCATGCACGCTGTT 552

Qy 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120

Db 553 TCTGTGGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAGAAGATGCC 612

Qy 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140

Db 613 AGTCTGGTGATCCAGCAGGCTTTGTGGCAGTGAATACCTACATGAGAATGGCATCGTC 672

Qy 141 HisArgAspLeuLys-----ProGluAsnLeuLeuTyrLeuThr 153

Db 673 CACAGAGACTTAAAGGTGTCAAGCGGGAGTCTCTGGCCCCGAAACCTGCTTTACCTTACC 732

Qy 154 ProGluGluAsnSerLysIleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsn 173

Db 733 CCTGAAGAGAACTCTAAGATCATGATCATCTACTGTTGTTGTTCTGTCCAGATGGAACAGAT 792

```
Qy 174 GlyIleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGluValIleuAlaGln 193
Db 793 GGCATCATGTCCACTGCCCTGTGGAGCCCGCAGATAGTGCGTCCAGAAAGTGTGCCCGAG 852
Qy 194 LysProTyrSerLysAlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeu 213
Db 853 AAACCCCTACAGCAAGCTGTGGATTGCTGCTCATCGCGCTCATCACCTACATATTGCTC 912
Qy 214 CysGlyTyrProProPheTyrGluThrGluSerLysLeuPheGluLysIleLysGlu 233
Db 913 TGTGGATACCCCATCTCATGAAGAACGGAGCTTAGCTTTTCGAGAAGATCAAGGAG 972
Qy 234 GlyTyrTyrGluPheGluSerProPheTrpAspIleSerGluSerAlaLysAspPhe 253
Db 973 GGCTACTATGAGTTTGAGTCTCCATCTCGGATGACATTCTTGAGTCAGCCAAAGGACTTT 1032
Qy 254 IleCysHisLeuLeuGluLysAspProAsnGluAtrGlyTyrThrCysGluLysAlaLeuSer 273
Db 1033 ATTTGCCACTTGCTTGAGAGGATCGGAACGAGCGGTACCTGTGAGAAGGCCCTTGAGT 1092
Qy 274 HisProTyrIleAspGlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeu 293
Db 1093 CATCCCTGGATTACAGGAACACAGCCCTCCACCGGGACATCTACCCATCAGTCAGCCTC 1152
Qy 294 GlnIleGlnLysAsnPheAlaLysSerLysTyrArgGlnAlaPheAsnAlaAlaVal 313
Db 1153 CAGATCCAGAAAGAACTTTGCTAGAGCAAGTGGAGCAAGCCCTCAACGCGCAGCAGCTGT 1212
Qy 314 ValHisHisMetArgLysLeuHisMetAsnLeuHisSerProGlyValArgProGluVal 333
Db 1213 GTGCACCATATGAGGAAGTACATCAATCAACCTGCACAGCCCGGGCGTCCGCCAGAGTG 1272
Qy 334 GluAsnArgProProGluThrGlnAlaSerGluThrSerArgProSerSerProGluIle 353
Db 1273 GAGAAACAGGCGCTGAAACTCAAGCTCAGAACTCTAGACCCAGCTCCCTCGAGATC 1332
Qy 354 ThrIleThrGluAlaProValLeuAspHisSerValAlaLeuProAlaLeuThrGlnLeu 373
Db 1333 ACCATCACCGAGGACCTGTCTCGACACAGTGTAGCACTCCCTGCGCTGACCCCAATTA 1392
Qy 374 ProCysGlnHisGlyArgArgProThrAlaProGlyGlyArgSerLysAsnCysLeuVal 393
Db 1393 CCCTGCCAGCATGGCGCGGCCACTGCCCTGTGTGGAGTCCCTCAACTGCCTGTGTC 1452
Qy 394 AsnGlySerLeuHisIleSerSerSerLeuValProMetHisGlnGlySerLeuAlaAla 413
Db 1453 AATGGCTCCCTCCACATCAGCAGCAGCTGGTGGTCCCATCATCAGGGGTCCCTGGCGGCC 1512
Qy 414 GlyProCysGlyCysSerSerCysLeuAsnIleGlySerLysGlyLysSerSerTyr 433
Db 1513 GGGCCCTGTGCTGTCTCCAGCTGCCGTGAACATTTGGAGCAAGGAAGTCCCTCTTAC 1572
Qy 434 CysSerGluProThrLeuLeuLysLysAlaAsnLysLysGlnAsnPhelLysSerGluVal 453
Db 1573 TGCTCTGAGCCCACTCTCTCAAAAAGGGCAACAAAAACAGAACTTCAAGTCGGAGGTC 1632
Qy 454 MetValProValLysAlaSerGlySerSerHisCysArgAlaGlyGlnThrGlyValCys 473
Db 1633 ATGTACACAGTTAAAGCAGTGGCAGCTCCCTCCCTGCGGGCAGGCGAGACTGGAGTCTGT 1692
Qy 474 LeuIleMet 476
Db 1693 CTCATTATG 1701
RESULT 4
AAS31014
ID AAS31014 standard; cDNA; 2689 BP.
AC AAS31014;
XX AAS31014;
XX 04-DEC-2001 (first entry)
XX
```

```
DE
XX
KW Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;
KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;
KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;
KW respiratory disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200162927-A2.
XX
XX 30-AUG-2001.
XX
XX 21-FEB-2001; 2001WO-US06059.
XX
XX 24-FEB-2000; 2000US-0184693.
XX
XX 24-FEB-2000; 2000US-0184697.
XX
XX 24-FEB-2000; 2000US-0184698.
XX
XX 24-FEB-2000; 2000US-0184768.
XX
XX 24-FEB-2000; 2000US-0184769.
XX
XX 24-FEB-2000; 2000US-0184770.
XX
XX 24-FEB-2000; 2000US-0184771.
XX
XX 24-FEB-2000; 2000US-0184772.
XX
XX 24-FEB-2000; 2000US-0184773.
XX
XX 24-FEB-2000; 2000US-0184774.
XX
XX 24-FEB-2000; 2000US-0184776.
XX
XX 24-FEB-2000; 2000US-0184777.
XX
XX 24-FEB-2000; 2000US-0184797.
XX
XX 24-FEB-2000; 2000US-0184813.
XX
XX 24-FEB-2000; 2000US-0184837.
XX
XX 24-FEB-2000; 2000US-0184841.
XX
XX 24-FEB-2000; 2000US-0185213.
XX
XX 24-FEB-2000; 2000US-0185216.
XX
XX 12-MAY-2000; 2000US-0203785.
XX
XX 15-MAY-2000; 2000US-0204226.
XX
XX 16-MAY-2000; 2000US-0204525.
XX
XX 16-MAY-2000; 2000US-0204821.
XX
XX 16-MAY-2000; 2000US-0204908.
XX
XX 16-MAY-2000; 2000US-0205232.
XX
XX 17-MAY-2000; 2000US-0204815.
XX
XX 17-MAY-2000; 2000US-0204863.
XX
XX 17-MAY-2000; 2000US-0205221.
XX
XX 17-MAY-2000; 2000US-0205285.
XX
XX 17-MAY-2000; 2000US-0205286.
XX
XX 17-MAY-2000; 2000US-0205287.
XX
XX 17-MAY-2000; 2000US-0205323.
XX
XX 17-MAY-2000; 2000US-0205324.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;
XX Chen A, D'Sa SA, Anshey S, Dahl CR, Dam TC, Daniels SE;
XX Dufour GE, Flores V, Fong WT, Greenawalt LB, Hillman JL, Jones AL;
XX Liu TF, Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Daffo A;
XX Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;
XX Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;
XX
XX WPI; 2001-502867/55.
XX
XX P-PSDB; AAU19443.
XX
XX Polynucleotides encoding diagnostic and therapeutic proteins, e.g.
XX enzymes, hormones and receptors, useful in diagnostics and therapeutics
XX
XX
XX Claim 1; Page 310-311; 522pp; English.
XX
XX The invention relates to polynucleotides (I) encoding diagnostic and
XX therapeutic (DITHP) polypeptides (II), which include e.g. enzymes,
XX and proteins involved in growth and development and receptors. (I) and
XX (II) may be used in the prevention, diagnosis and treatment of diseases
XX associated with inappropriate DITHP expression. For example, (I) and
XX (II) may be used to treat disorders associated with decreased polypeptide
XX expression by rectifying mutations or deletions in a patient's genome,
```

CC that affect the activity of the DITHPs, by expressing inactive proteins  
 CC or supplementing the patient's own production of them. (I) and (II)  
 CC may be used to treat diseases, for example, cell proliferative disorder,  
 CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,  
 CC leukaemia, autoimmune disorders, and respiratory disorders. Additionally,  
 CC (I) may be used to produce the DITHPs, by inserting the nucleic acids  
 CC into a host cell and culturing the cell to express the protein. (I) and  
 CC its complementary sequences may also be used as DNA probes in diagnostic  
 CC assays to detect and quantitate the presence of similar nucleic acids in  
 CC samples, and therefore which patients may be in need of restorative  
 CC therapy. (II) may also be used as antigens in the production of  
 CC antibodies against DITHPs and in assays to identify modulators of DITHP  
 CC expression and activity. The anti-DITHP antibodies and antagonists may  
 CC also be used to down regulate expression and activity. The anti-DITHP  
 CC antibodies may also be used as diagnostic agents for detecting the  
 CC presence of DITHPs in samples (e.g. by enzyme linked immunosorbant  
 CC assay (ELISA)). AAS30986-AAS31196 represent human diagnostic and  
 CC therapeutic (DITHP) polynucleotides of the invention.  
 XX  
 SQ Sequence 2689 BP; 642 A; 748 C; 678 G; 621 T; 0 other;

## Alignment Scores:

Pred. No.: 7,06e-190 Length: 2689  
 Score: 2327.50 Matches: 460  
 Percent Similarity: 95.86% Conservative: 3  
 Best Local Similarity: 95.24% Mismatches: 13  
 Query Match: 92.62% Indels: 7  
 DB: 22 Gaps: 1

US-09-960-643-2 (1-476) x AAS31014 (1-2689)

QY 1 MetGlyArgLysGluGluAspCysSerTrpLysLysGlnThrAsnIleArg 20  
 DB 237 ATGGTTCGAAGGAGAGAGTGAAGTCACTTCTTGGAGAAACAGACACCAACATCCGG 296  
 QY 21 LysThrPheIlePheKetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
 DB 297 AAAACCTTCAATTTTATGGAAGTGCCTGGGATCAGGAGCTTCTCAAGATTTTCTCGTG 356  
 QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60  
 DB 357 AAGCAAGACTGACTGGGAGCTCTTTGCTCTGAAGTGCATCAAGAATCACCCTGCCCTTC 416  
 QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80  
 DB 417 CGGACAGCAGCCTGGAGATGAGATTGCTGTGTTGAAAAGATCAAGCATGAAAACATT 476  
 QY 81 ValThrLeuGluAspIleTrpGluSerThrThrHisTrpLysValMetGlnLeuVal 100  
 DB 477 GTACCTTGGAGGACATCTATGAGACACACCCCTACTACTGTGTCATGAGCTTGTGT 536  
 QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTrpThrGluLysAspAla 120  
 DB 537 TCTGGTGGGAGCTCTTTGACCGGATCTCTGAGCGGGGTGTACACAGAGAAGATGCC 596  
 QY 121 SerLeuValIleGlnValLeuSerAlaValLysTrpLysLeuHisGluAsnGlyIleVal 140  
 DB 597 AGTCTGGTGATCCAGCAGGTCTTGTGCGCAGTGAATACATACATGAGAATGGCATCGTC 656  
 QY 141 HisArgAspLeuLysProGluAsnLeuLeuTrpLysLeuThrProGluAsnSerLysIle 160  
 DB 657 CACAGAGACTTAAGGCCGGAACCTGCTTTACCTTACCCCTGAGAGAACTCTAAGATC 716  
 QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180  
 DB 717 ATGATCACTGACTTGTCTGTCGAAGATGGAACAGATGGCATCATGCTCCACTGCCTGT 776  
 QY 181 GlyThrProGlyTrpValAlaProGluValLeuAlaGlnLysProTrpSerLysAlaVal 200  
 DB 777 GGGACCCCAAGGCTACGTGGCTCCAGAAAGTGTGGCCCAAGAAACCTTACAGCAAGGCTGTG 836  
 QY 201 AspCysTrpSerIleGlyValIleThrTrpIleLeuLeuCysGlyTrpProPheTrp 220

DB 837 GATTGCTGGTCCATCGGCGTCACTCACCTACATATTGCTGTGTGATACCCCACTTCTAT 896  
 QY 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTrpTrpGluPheGluSer 240  
 DB 897 GAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGGCTACTATGAGTTTGAGTCT 956  
 QY 241 ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260  
 DB 957 CCATTCCTGGGATGACATTTCTGAGTCAGCCAGGACTTTATTGGCCACTTGCCTGAGAAG 1016  
 QY 261 AspProAsnGluArgTrpThr-CysGluLysAlaLeuSerHisProTrpIleAspGlyAs 280  
 DB 1017 GATCCGAACGAGGGGTAAACACTGTGAGAAGGGCTTGAGTCATCCCTCGGATTGACGGTAA 1076  
 QY 280 nThrAlaLeuHisArgAspIleTrpProSerVal-SerLeuGlnIleGlnLysAsnPheA 300  
 DB 1077 CACAGCCCTCCACGGGACATCTACCCATCAGTCAGGCGTCCAGATCCAGAAACACTTTG 1136  
 QY 300 laLysSerLysTrpArgGlnAlaPheAsnAlaAlaAlaValVal-HisHisMetArgLys 319  
 DB 1137 CTAGAGCAAGTGGAGGCAAGCCTTCAACGACGAGCTGTGGTAGCACACAATGAGGAAG 1196  
 QY 320 LeuHisMetAsnLeuHisSer-ProGlyValArgProGluValGluAsnArgProProG 339  
 DB 1197 CTACACATGAACCTGCACAGCCCGGGCGTCCGCCAGAGGTGGAGAACAGGCGCGCTGA 1256  
 QY 339 uThrGlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaPr 359  
 DB 1257 AACTCAAGCCTCAGAAACCTCTAGACCCAGCTCCCTCTGAGATCACCATCACCAGGACCC 1316  
 QY 359 oValLeuAspHisSerValAlaLeuProAlaLeuThrGlnLeu---ProCysGlnHisG 378  
 DB 1317 TGTCTCGGACCACAGTGTAGCACTCCCTGCTCTGACCCCAATATACCTGCCAGCATCG 1376  
 QY 378 yArgArgProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuH 398  
 DB 1377 CCGCGGCGCCACTGCCCTTGGTGGCAGGTCCCTCAACTGCTGCTCAATGGCTCCCTCCA 1436  
 QY 398 sIleSerSerSerLeuValProMetHisGlnGlySerLeuAlaGlyProCysGlyCy 418  
 DB 1437 CATCAGCAGCAGCCTGGTGGCCATGCATAGGGGTCCCTGCGCGCGGCGCTGTGGCTG 1496  
 QY 418 sCysSer-SerCysLeuAsnIleGlySerLysGlyLysSerSerTrpCysSerGluPro 438  
 DB 1497 CTGCTCCAAGCTGCTGAACATTGGGAGCAAGAAAGTCTCTCTACTGCTGTGAGCCCA 1556  
 QY 438 hrLeuLeuLysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProVal 458  
 DB 1557 CACTCTCAAAAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTCTGTACCGGTTA 1616  
 QY 458 yAlaSerGly-SerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476  
 DB 1617 AAGCCAGTGGCATGCTTCCACTGCGGCGGAGGAGAGTCTGCTCATATTG 1673

## RESULT 5

AAD18817  
 ID AAD18817 standard; cDNA; 1736 BP.

XX AAD18817;

DT 18-DEC-2001 (first entry)

XX Human kinase (PKIN)-2 cDNA.

XX Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;  
 KW cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;  
 KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;  
 KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;  
 KW myasthenia gravis; cirrhosis; cataract; growth and development disorder;  
 KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;  
 KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;  
 KW asthma; obesity; restorative therapy; cytostatic; immunomodulatory;  
 KW antimicrobial; cardiovascular; antiinflammatory; vaccine; ss.

XX OS Homo sapiens.  
 XX Key Location/Qualifiers  
 PH CDS 159..1232  
 FT /\*tag= a  
 FT /product= "Human PKIN-2 protein"  
 FT sig\_peptide  
 FT mat\_peptide  
 FT 279..1229  
 FT /\*tag= b  
 FT /\*tag= c  
 FT /product= "Mature human PKIN-2 protein"

XX W0200181555-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12992.

XX 20-APR-2000; 2000US-199021P.

XX 28-APR-2000; 2000US-200226P.

XX 05-MAY-2000; 2000US-202339P.

XX 11-MAY-2000; 2000US-203505P.

XX 18-MAY-2000; 2000US-205564P.

XX 26-MAY-2000; 2000US-207739P.

XX 01-JUN-2000; 2000US-208795P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;  
 PI Bandman O, Lu DAM, Lal P, Burford N, Khan FA, Walia NK, Yao MG;  
 PI Patterson C, Burrill JB, Marcus GA, Zingler KA, Recipon SA, Lu Y;  
 PI Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;  
 PI Walsh RT, Ramkumar J, Borowsky ML, Au-young J, Hillman JL;  
 PI Gururajan R;

XX WPI: 2001-611740/70.

XX P-PSDB; AAEL1768.

XX Human kinases and nucleic acids, useful for preventing diagnosing and  
 PT treating cancers, inflammation and immune disorders -

XX Claim 5: Page 151-152; 166pp; English.

XX The present invention relates to human kinases (PKIN) and the nucleic  
 CC acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is  
 CC used in the prevention, diagnosis and treatment of diseases cancers,  
 CC adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,  
 CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,  
 CC gout, microbial infections, cardiovascular disease and/or inflammation,  
 CC myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial  
 CC infarction, cataract, growth and development disorder, seizure disorder,  
 CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage  
 CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.  
 CC PKIN may be used to treat disorders associated with decreased PKIN  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of PKIN by expressing inactive proteins or to  
 CC supplement the patients own production of PKIN. PKIN nucleic acids may be  
 CC used to produce the PKIN polypeptide, by inserting the nucleic acids into  
 CC a host cell and culturing the cell to express the protein. PKIN nucleic  
 CC acid and its complementary sequences may also be used as DNA probes in  
 CC diagnostic assays to detect and quantitate the presence of similar  
 CC nucleic acid sequences in samples and therefore which patients may be in  
 CC need of restorative therapy. The present sequence is human PKIN-2 cDNA.

XX SQ Sequence 1736 BP; 463 A; 456 C; 466 G; 351 T; 0 other;

XX Alignment Scores:

Pred. No.: 1.22e-97 Length: 1736  
 Score: 1250.50 Matches: 272  
 Percent Similarity: 62.85% Conservative: 63  
 Best Local Similarity: 51.03% Mismatches: 110  
 Query Match: 49.76% Indels: 90

DB: 22 Gaps: 11  
 US-09-960-643-2 (1-476) x AAD18817 (1-1736)  
 Qy 1 MetGlyArgLysGluGluAspAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20  
 Db 159 ATGCCCGGGAGACGGCAGAGCAGCTCTCTCTGGAAAAAGCAAGCTGAAGCATCAAG 218  
 Qy 21 LysThrPheIlePheMetGluValLysGlySerGlyAlaPheSerGluValPheLeuVal 40  
 Db 219 AAGATCTTCGAGTTCAAAAGAGACCTCGAAGCCGGGCCCTTTCCGAAGTGGTTAGCT 278  
 Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59  
 Db 279 GAAGAGAAGCAACTGGCAAGCTCTTGTCTGTGAAGTGTATCCTTAAGAAGCGCTGAAG 338  
 Qy 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysLysLysGluAsn 79  
 Db 339 GGCAAGGAAAGCAGCATAGAGATGAGATAGCGCTCTCGAAGATTAAGCATGAAAT 398  
 Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99  
 Db 399 ATTGTTGCCCTGGGAAGACATTTATGAAGGCCCAATCACCTGACTTGGTCATGCAGCTG 458  
 Qy 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119  
 Db 459 GTGTCCGGTGGAGAGCTGTTGACCGGATAGTGGAGAGGGGTTTATACAGAGAGGAT 518  
 Qy 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139  
 Db 519 GCCAGCATCTGTATGCCCAAGTCTTGGACGGCTCTACTATCTCCACAGAATGGGCATC 578  
 Qy 140 ValHisArgAspLeuLysProGluAsnLeuTyrLeuThrProGluGluAsnSerLys 159  
 Db 579 GTCACAGAGACCTCAAGCCCGAAATCTCTTGTACTACAGTCAAGTCAAGAGGAGTCCAA 638  
 Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178  
 Db 639 ATAATGATCATGACTTGGATTGTTCAAAAATGGAGGGCAAGAGATGTGATGCCACT 698  
 Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198  
 Db 699 GCCTGTGGAATCCAGCTATGCTGCTCTGAAAGTCTCGCCAGAAACCTTACAGCAA 758  
 Qy 199 AlaValAspCysTpsSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218  
 Db 759 GCCGTTCACTGCTGGTCCATCGGAGTGATGCTCTACATCTTGTCTCGCGCTACCTCCT 818  
 Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238  
 Db 819 TTTTATGATGAAATGACTCCAAGCTCTTTGAGCAGATCTCAAGCGGGAATATGAGTTT 878  
 Qy 239 GluSerProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258  
 Db 879 GACTCTCCCTACTGGGATGACATCTCGACTCTGCAAAAGACTTCATTCGGAACCTGATG 938  
 Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278  
 Db 939 GAGAAGGACCCGATAAAAGATACACGTGTGAGCAGCAGCTCGGCACCCATCGATCGCT 998  
 Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298  
 Db 999 GGTGACACAGCCCTCAACAAAAACATCCAGAGTCCGTACAGCCCGCAGATCCGGAATAAC 1058  
 Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArg 318  
 Db 1059 TTGCCCCAAGAGCAAAATGGAGACAAGCATTTAATGCCACGCCCTGTGAGACATATGAGA 1118  
 Qy 319 LysLeuHisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArg----- 336  
 Db 1119 AAACATACACCTCGCAGCAGCCTCGGACAGTTCAAATGCAAGTGTTCGACGACCTCAGT 1178  
 Qy 337 ---ProProGluThrGlnAlaSerGluThrSerArgProSerProGluIleThrIle 355



```
Db 541 GCCTGTGGAACCTCCAGGCTATGTGCTCGCTCGCTGAAGTCTCTGCCCGCAGAAACCTTACAGCAA 600
Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
Db 601 GCCGTTGACCTGCTGGTCAATCGGAGTGATGCTACATCTTGCTCTGCGGCTACCTCCT 660
Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238
Db 661 TTTTATGATGAAATGACTCCAAAGCTCTTTGACGACATCCTCAAGCGGGAATATGAGTTT 720
Qy 239 GluSerProPheTrpAspPheSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
Db 721 GACTCTCCCTACTGGGATGACATCTCGGACTCTGCAAAAGACTTCAATTCGGAACCTGATG 780
Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278
Db 781 GAGAAGACCCGCAATAAAGATACAGTGTGTGACGAGCAGCTCGGCACCCATGGATCGCT 840
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGluIleGlnLysAsn 298
Db 841 GGTGACACACCCCTCAACAAAAACATCCAGAGTCCGTGAGCGCCCGCAGATCCGGAAC 900
Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318
Db 901 TTTGCCAAGAGCAATGGAGACAAGCATTTAATGCCACCGCCGCTGCTGAGACATATGAGA 960
Qy 319 LysLeuHisMet 322
Db 961 AAACACTACACTC 972

RESULT 7
AAH25119
ID AAH25119 standard; DNA; 1074 BP.
XX
AC AAH25119;
XX
DT 22-AUG-2001 (first entry)
XX
DE Nucleotide sequence of a human kinase polypeptide.
XX
KW Human; kinase; human disease; human disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1074
FT /tag= a
FT /transl_except= "(pos: 838..843, aa: Ala)"
FT /product= "kinase"
XX
XX W0200142435-A2.
XX
XX 14-JUN-2001.
XX
XX 07-DEC-2000; 2000WO-US33240.
XX
XX 07-DEC-1999; 99US-0169428.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Donoho G, Scoville J, Turner CA, Friedrich G, zambrowicz B;
XX Abuin A, Sands AT;
XX
XX WPI; 2001-381667/40.
XX P-PSDB; AAB84360.
XX
XX Novel isolated human kinase polynucleotide that shares structural
XX similarity with animal kinases including calcium/calmodulin-dependent
XX protein kinases and serine/threonine protein kinases, useful in
XX therapeutics -
XX
XX Disclosure; Page 30-31; 32pp; English.
XX
```

CC The present sequence encodes a kinase polypeptide. The kinase  
CC polynucleotides and polypeptides are useful in therapeutic, diagnostic  
CC and pharmacogenetic applications. They are useful for the detection of  
CC mutant kinases, or inappropriately expressed kinases for the diagnosis  
CC of a disease or disorder. They are useful for screening for drugs (or  
CC high throughput screening of combinatorial libraries) effective in the  
CC treatment of symptomatic or phenotypic manifestations of that disease  
CC or disorder. The polynucleotide sequence is useful as a source of  
CC probes and primers, which can be used to screen libraries, isolate  
CC clones, and prepare cloning and sequencing templates.

SQ Sequence 1074 BP; 313 A; 258 C; 276 G; 227 T; 0 other;

#### Alignment Scores:

Pred. No.: 1.5e-97 Length: 1074  
Score: 1246.00 Matches: 237  
Percent Similarity: 85.49% Conservative: 40  
Best Local Similarity: 73.15% Mismatches: 45  
Query Match: 49.58% Indels: 2  
Db: 22 Gaps: 2

US-09-960-643-2 (1-476) x AAH25119 (1-1074)

```
Qy 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
Db 1 ATGGCCCGGAGAACGGCGAGAGCAGCTCTCTCTGGAAAAAGCAAGCTGAAGCATCAAG 60
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 61 AAGATCTTCGAGTTCAAAGAGACCCCTCGGAACCGGGGCCCTTTCGGAAGTGGTTTAGCT 120
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59
Db 121 GAAGAGAGAGGCAACTGGCAAGCTCTTTGCTGTGAAGTGTATCCCTAAGAAAGCGCTGAAG 180
Qy 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
Db 181 GCGAAGGAAACACACATAGAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 240
Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99
Db 241 ATTGTGCGCTGGAAGACATTTATGAAAGCCAAATCACCTGTACTTGGTTCATGACGCTG 300
Qy 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119
Db 301 GTGTCGCGTGGAGAGCTGTTTGACCGGATAGTGAGAGGGGGTTTTATACAGAGAGGAT 360
Qy 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139
Db 361 GCCAGCCTCTGATCCGCAAGTCTTGGACCGCTGTACTATCTCCACAGAAATGGGCATC 420
Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159
Db 421 GTCCACAGAGACCTCAAGCCGGAATCTCTTGCTACTACAGTCAAGATGAGGAGTCCAAA 480
Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178
Db 481 ATAATGATCATGTGACTTTGGATTGTCAAAATGAGGAGGCAAGAGATGTGATGTCCACT 540
Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198
Db 541 GCCTGTGAACTCCAGGCTATGCTCGCTCGTGAAGTCTCGCCCAAGAAACCTTACAGCAA 600
Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuLysGlyTyrProPro 218
Db 601 GCCGTTGACCTGCTGGTCAATCGGAGTGATGCTACATCTTGCTCTGCGGCTACCTCCT 660
Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238
Db 661 TTTTATGATGAAATGACTCCAAAGCTCTTTGAGCAGATCTCAAGCGGGAATATGAGTTT 720
Qy 239 GluSerProPheTrpAspPheSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
Db 721 GACTCTCCCTACTGGGATGACATCTCGGACTCTGCAAAAGACTTCAATTCGGAACCTGATG
```

Db 721 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAGACTTCAATTCGGAACTGATG 780

Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278

Db 781 GAGAAGAGCCCGAATAAAGATACAGTGTGAGCAGGACGTCGGCACCATGGATCGCT 840

Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298

Db 841 GGTGACACAGCCCTCAACAAAAACATCCACGAGTCGTCAGCGCCAGATCGGAAAAAC 900

Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValAlaValHisMetArg 318

Db 901 TTGTCCCAAGCAAAATGGAGACAAGCATTTAATGCCACGCGCTGCTGAGACATATGAGA 960

Qy 319 LysLeuHisMet 322

Db 961 AACTACACTC 972

RESULT 8

AAH25118

ID AAH25118 standard; DNA; 1158 BP.

XX

AC AAH25118;

XX

DT 22-AUG-2001 (first entry)

XX

DE Nucleotide sequence of a human kinase polypeptide.

XX

KW Human; kinase; human disease; human disorder; ss.

XX

OS Homo sapiens.

XX

FH Key

FT 1..1158

FT CDS

FT /\*tag= a

FT /product= "kinase"

XX

PN W0200142435-A2.

XX

PD 14-JUN-2001.

XX

PF 07-DEC-2000; 2000WO-US33240.

XX

PR 07-DEC-1999; 99US-0169428.

XX

PA (LEXI-) LEXICON GENETICS INC.

XX

PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;

PI Abuin A, Sands AT;

XX

DR WPI; 2001-381667/40.

DR P-PSDB; AAB84359.

XX

PT Novel isolated human kinase polynucleotide that shares structural

PT similarity with animal kinases including calcium/calmodulin-dependent

PT protein kinases and serine/threonine protein kinases, useful in

PT therapeutics -

XX

PS Claim 1; Page 29; 32pp; English.

XX

CC The present sequence encodes a kinase polypeptide. The kinase

CC polynucleotides and polypeptides are useful in therapeutic, diagnostic

CC and pharmacogenetic applications. They are useful for the detection of

CC mutant kinases, or inappropriately expressed kinases for the diagnosis

CC of a disease or disorder. They are useful for screening for drugs (or

CC high throughput screening of combinatorial libraries) effective in the

CC treatment of symptomatic or phenotypic manifestations of that disease

CC or disorder. The polynucleotide sequence is useful as a source of

CC probes and primers, which can be used to screen libraries, isolate

CC clones, and prepare cloning and sequencing templates.

XX

Sequence 1158 BP; 323 A; 283 C; 304 G; 248 T; 0 other;

Alignment Scores:

Pred. No.: 1.67e-97 Length: 1158

Score: 1246.00 Matches: 237

Percent Similarity: 85.49% Conservative: 40

Best Local Similarity: 73.15% Mismatches: 45

Query Match: 49.58% Indels: 2

DB: 22 Gaps: 2

US-09-960-643-2 (1-476) x AAH25118 (1-1158)

Qy 1 MetGlyArgLysGluLysAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20

Db 1 ATGCCCGGGAGACGCGGAGAGCAGCTCTCTCCCTGGAAAAAGCAAGCTGAAGCATCAAG 60

Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40

Db 61 AAGATCTTCGAGTTCAAAGAGACCTCGGAACCGGGGCTTTTCCGAAGTGTGTTAGCT 120

Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59

Db 121 GAAGAGAGGCAACTGGCAAGCTCTTTGCTGTAAGTGTATCCCTAAGAAGCGCGCTGAAG 180

Qy 60 PheArgAspSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79

Db 181 GCAAGAGAAACAGCATAGAGATGAGATAGCCGCTCTGAGAAAGATTAACCATGAAAT 240

Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99

Db 241 ATTGTTGCCCTTGGAGACATTTATGAAGCCCAATCACCTGTACTTGTGGTTCATGAGCTG 300

Qy 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119

Db 301 GTGTCCGCTGGAGAGCTGTTGACCGGATAGTGGAAAGGGGTTTATACAGAGAAGAT 360

Qy 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139

Db 361 GCCAGCACTCTGATCCGCCAAGTCTTGGACCGCTGTACTATCTCCACAGAATGGCATC 420

Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159

Db 421 GTCCACAGAGACCTCAAGCCGGAATCTCTGTACTACAGTCAAGATGAGAGATGCCAAA 480

Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178

Db 481 ATAATGATCAGTGACTTTGGATTTGCAAAATGGAGGCAAGAGAGATGTGATGCCACT 540

Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198

Db 541 GCCTGTGGAACCTCCAGGCTATGTCGCTCTGAAGTCTCTGCCAGAAAACCTTACAGCAA 600

Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuLysGlyTyrProPro 218

Db 601 GCGGTGTGACTGCTGCTCCATCGAGTGTGCTACATCTTGTCTGCGGGTACCCCTCT 660

Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238

Db 661 TTTTATGATGAAAATGACTCCCAAGCTCTTTGAGCAGATCTCAAGGCGGATATGAGTTT 720

Qy 239 GluSerProPheTrpAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258

Db 721 GACTCTCCCTACTGGATGACATCTCCGACTCTGCAAAAGACTTTCATTTCGGAACCTGATG 780

Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278

Db 781 GAGAAGAGCCCGAATAAAGATACACGTTGTGAGCAGGAGCTCGGACCCCATGGATCGCT 840

Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298

Db 841 GTGTACACAGCCCTCAACAAAAACATCCAGGAGTCCGTCAGCGCCAGATCGGAAAAAC 900

Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValAlaValHisMetArg 318

Db 901 TTGTCCCAAGCAAAATGGAGACAAGCATTTAATGCCACGCGCTGCTGAGACATATGAGA 960

QY 319 LysLeuHisMet 322  
 Db 961 AACTACACCTC 972  
 RESULT 9  
 ID ABL60905  
 XX ABL60905 standard; cDNA; 1565 BP.  
 AC ABL60905;  
 XX  
 DT 23-SEP-2002 (first entry)  
 DE Human Cam kinase I 39.05 polypeptide encoding cDNA.  
 XX  
 KW Human; Cam kinase I 39.05; nervous disease; arrhythmia; tumour;  
 KW gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 18..1085  
 FT /\*tag= a  
 FT /product= "human Cam kinase I 39.05"  
 XX  
 PN CN1333357-A.  
 XX  
 PD 30-JAN-2002.  
 XX  
 PF 07-JUL-2000; 2000CN-0119405.  
 XX  
 PR 07-JUL-2000; 2000CN-0119405.  
 XX  
 PA (SHAN-) SHANGHAI BIODOOR GENE DEV CO LTD.  
 XX  
 PI Mao Y, Xie Y;  
 XX  
 DR WPI; 2002-305611/35.  
 DR P-PSDB; ABB08178.  
 XX  
 PT Novel polypeptide human Cam kinase I 39.05 and polynucleotide for  
 PT encoding said polypeptide -  
 XX  
 PS Claim 6; Page 26-27 (disclosure); 35pp; Chinese.  
 XX  
 CC The present invention discloses a novel polypeptide-human Cam kinase I  
 CC 39.05, polynucleotide for coding the polypeptide and method for producing  
 CC this polypeptide by DNA recombination technology. The polypeptide is  
 CC useful for treating diseases such as nervous disease, arrhythmia, tumour  
 CC and growth development disturbance disease. The present sequence  
 CC represents the human Cam kinase I 39.05 polypeptide encoding cDNA.  
 XX  
 SQ Sequence 1565 BP; 408 A; 394 C; 410 G; 353 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 2.56e-97 Length: 1565  
 Score: 1246.00 Matches: 237  
 Percent Similarity: 85.49% Conservative: 40  
 Best Local Similarity: 73.15% Mismatches: 45  
 Query Match: 49.58% Indels: 2  
 DB: 24 Gaps: 2  
 US-09-960-643-2 (1-476) x ABL60905 (1-1565)  
 QY 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20  
 Db 18 ATGGCCCGGAGAACGGCAGCAGCTCCTCGAAGAAAGCAAGCTGAAGACATCAAG 77  
 QY 21 LysThrPheIlePheMetGluValLysGlySerGlyAlaPheSerGluValPheLeuVal 40  
 Db 78 AGATCTTCGAGTTCAAGAGACCTCGAACCAGGGGCTTTCCGAGTGGTTTAGCT 137  
 QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59

Db 138 GAAGAGAGGCAACTGGCAAGCTCTTTGCTGTGAAGTGTATCCCTAAGAGGCGCTGAAG 197  
 QY 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79  
 Db 198 GGCAAGGAAGCAGCATAGAGATGAGATAGCGCTCTGAGAAGATTAAAGCATGAAAT 257  
 QY 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99  
 Db 258 ATTGTTGCCCTGGGAAGACATTTATGAAAGCCCAATACACCTGTACTTGGTCATGCAGTG 317  
 QY 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119  
 Db 318 GTGTCCGGTGGAGAGCTGTTTGACCGGATAGTGAGAGGGGGTTTATACAGAGAAGGAT 377  
 QY 120 AlaSerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139  
 Db 378 GCCAGCACTCTGATCCGCCAAGTCTTGGAGCGCGTGTACTATCTCCACAGAATGGCATC 437  
 QY 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159  
 Db 438 GTCCACAGAGACCTCAAGCCGAAATCTCTGTACTACAGTCAAGATGAGGATCCAAA 497  
 QY 160 IleMetIleThrAspPheGlyLeuSerLysMetGluLysAsnGly---IleMetSerThr 178  
 Db 498 ATATGATCATGTGACTTTGGATTGTCAAAATGGAGGGCAAGGAGATGTGTCCACT 557  
 QY 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198  
 Db 558 GCCTGTGGAACCTCCAGCTATGCTCGCTCGTGAAGTCTCGCCACAGAACTTACACAAA 617  
 QY 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218  
 Db 618 GCGTGTGACTGCTGGTCCATCGGAGTGTGCTGTACTATCTGCTCTGCGCTACCTCCT 677  
 QY 219 PheTyrGluGluThrGluSerLysLeuPheGluLysLysGluGlyTyrTyrGluPhe 238  
 Db 678 TTTTATGATGAAAATGACTCCAAAGCTCTTGGACAGATCTCAAGCGGGAATATGAGTTT 737  
 QY 239 GluSerProPheTrpAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258  
 Db 738 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAGACTTCAATCGGAACCTGATG 797  
 QY 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278  
 Db 798 GAGAAGGACCCGAATAAAGATACACGTGTGAGCAGGCGACTCGGCACCCATGGATCGCT 857  
 QY 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298  
 Db 858 GGTGACACAGCCCTCAACAAAAACATCCACGAGTCCGTGAGCCGCCAGATCCGGAANAAC 917  
 QY 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318  
 Db 918 TTTGCCAAGAGCAAAATGGAGACAAGCATTTAATGCCACGCCGCTCGTGAGACATATGAGA 977  
 QY 319 LysLeuHisMet 322  
 Db 978 AACTACACCTC 989  
 RESULT 10  
 ID AAD18826  
 XX AAD18826 standard; cDNA; 1661 BP.  
 AC AAD18826;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Human kinase (PKIN)-11 cDNA.  
 XX  
 KW Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;  
 KW cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;  
 KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;  
 KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;

KW myasthenia gravis; cirrhosis; cataract; growth and development disorder;  
KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;  
KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;  
KW obesity; restorative therapy; immunomodulatory; vaccine; cardiovascular;  
KW antimicrobial; cytostatic; antiinflammatory; asthma; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 159..1226

FT /\*tag= a

FT /product= "Human PKIN-11 protein"

XX WO200181555-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12992.

XX 20-APR-2000; 2000US-199021P.

PR 28-APR-2000; 2000US-200226P.

PR 05-MAY-2000; 2000US-202339P.

PR 11-MAY-2000; 2000US-203505P.

PR 18-MAY-2000; 2000US-205564P.

PR 26-MAY-2000; 2000US-207739P.

PR 01-JUN-2000; 2000US-208795P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;

PI Bandman O, Lu DAM, Lal P, Burford N, Khan FA, Walla NK, Yao MG;

PI Patterson C, Burrill JD, Marcus GA, Zingler KA, Recipion SA, Lu Y;

PI Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;

PI Walsh RT, Ramkumar J, Borowsky ML, Au-young J, Hillman JL;

PI Gururajan R;

XX WPI; 2001-611740/70.

DR P-PSDB; AAE11777.

XX Human kinases and nucleic acids, useful for preventing diagnosing and

PT treating cancers, inflammation and immune disorders -

XX Claim 5; Page 159; 166pp; English.

XX The present invention relates to human kinases (PKIN) and the nucleic  
CC acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is  
CC used in the prevention, diagnosis and treatment of diseases cancers,  
CC adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,  
CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,  
CC gout, microbial infections, cardiovascular disease and/or inflammation,  
CC myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial  
CC infarction, cataract, growth and development disorder, seizure disorder,  
CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage  
CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.  
CC PKIN may be used to treat disorders associated with decreased PKIN  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of PKIN by expressing inactive proteins or to  
CC supplement the patients own production of PKIN. PKIN nucleic acids may be  
CC used to produce the PKIN polypeptide, by inserting the nucleic acids into  
CC a host cell and culturing the cell to express the protein. PKIN nucleic  
CC acid and its complementary sequences may also be used as DNA probes in  
CC diagnostic assays to detect and quantitate the presence of similar  
CC nucleic acid sequences in samples and therefore which patients may be  
CC in need of restorative therapy. The present sequence is human PKIN-11  
CC cDNA.

XX SQ Sequence 1661 BP; 408 A; 452 C; 448 G; 353 T; 0 other;

XX Alignment Scores:

XX Pred. No.: 2,79e-97 Length: 1661

XX Score: 1246.00 Matches: 237

XX Percent Similarity: 85.49% Conservativity: 40

XX Best Local Similarity: 73.15% Mismatches: 45

Query Match: 49.58% Indels: 2  
DB: 22 Gaps: 2  
US-09-960-643-2 (1-476) x AAD18926 (1-1661)  
QY 1 MetGlyArgLysGluGluAspAspCysSerTrpLysGlnThrThrAsnIleArg 20  
DB 159 ATGCCCGGGAACGCGGAGAGCAGCTCCTCTCGTAAAGAACGAGCATCAAG 218  
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSercluValPheLeuVal 40  
DB 219 AAGATCTTCAGTTCAAAGAGAGCCCTCGGAACGGGGCCCTTTCCGAAGTGGTTTATAGCT 278  
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59  
DB 279 GAAGAGAAGCAACTGCAAGAGCTCTTTGCTGTCAAGTGTATCCCTAAGAAGCGCTGAAG 338  
QY 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79  
DB 339 GCGAAGGAAAGCAGCATAGAGATGAGCGCTCTCGAAGAGATTAAAGCATGAAAT 398  
QY 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99  
DB 399 ATTGTTGCCCTGGAAGACATTTATGAAGCCCAATCACCTGTACTTGTCTATGTCAGCTG 458  
QY 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119  
DB 459 GTCTCCGGTGGAGAGCTGTTTGACCGGATAGTGGAGAGGGGTTTTATACAGAGAAGAT 518  
QY 120 AlaSerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139  
DB 519 GCCAGCAGCTGTGATCGGCCCAAGTCTTTGGACGCGCTACTATCTCCACAGAATGGGCATC 578  
QY 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159  
DB 579 GTCCACAGAGACCTCAAGCCCGAAATCTCTGTACTACGTCAAGATGAGGAGTCCAAA 638  
QY 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178  
DB 639 ATAATGATCAGTCAGCTTTGGATTGTCAAAATGGAGGGCAAGGAGATGTGATGTCCACT 698  
QY 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198  
DB 699 GCCTGTGGAACTCCAGGCTATGTCGCTCTGAAAGTCTCGCCAGAAACCTTACAGCAA 758  
QY 199 AlavalAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218  
DB 759 GCCTGTGACTGCTGCTCCATCGGAGTGATGCTACATCTTGCTCTCGGCTACCCCTCCT 818  
QY 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238  
DB 819 TTTTATGATGAAATGACTCCCAAGCTCTTTTGAGCAGATCCTCAAGCGGGAATATGAGTTT 878  
QY 239 GluSerProPheThrAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258  
DB 879 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAGACTTCATTCGGAACTGATG 938  
QY 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278  
DB 939 GAGAAGGACCGGAATAAAGATACACGTGTGAGCAGGAGCTCGGCACCCATGGATCGCT 998  
QY 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298  
DB 999 GGTGACACAGCCCTCAACAAAACATCCACGAGTCCGTTCAGCGCCGACAGATCCGGAAC 1058  
QY 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArg 318  
DB 1059 TTGTCCCAAGAGCAAAATGGAGACAGCATTTAATGCCCGCGCTCGTGAGACATATGAGA 1118  
QY 319 LysLeuHisMet 322  
DB 1119 AAATAACAGCTC 1130

[illegible]

QY 385 -----GlyGlyArgSerLeuAsnCysLeuValAsn 394  
|||||  
Db 1485 TCTACTAAATATAAATAGCCGGTGTGGCGGACACCTGTAACTGCTACTT 1544  
|||  
QY 395 GlySerLeuHisIleSerSerLeuValPromethHisGlnGlySerLeuAlaAlaGly 414  
|||  
Db 1545 GGGAGGCTGAGGCA-GGAGAAATCGCTTGAGCCC-----GGA 1579  
|||||  
QY 415 ProCysGlyCysCysSerSerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCys 434  
|||||  
Db 1580 AGCGGAGGTGC-----AATGAGCTGAGATCACCACCTGC 1615  
|||||  
QY 435 Ser-GluPro-----ThrLeuLeuLysLysAlaAsnLysLysGln 447  
|||  
Db 1616 ACTTCAGCCTGGGTGACAGATTGAGACTCTCTCAAAAAAATAAAAAA 1670  
|||  
RESULT 12  
AAI60424  
ID AAI60424 standard; cdna; 2164 BP.  
XX AC AAI60424;  
XX AC AAI60424;  
XX DT 22-OCT-2001 (first entry)  
XX DE Human polynucleotide SEQ ID NO 4413.  
XX KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX OS Homo sapiens.  
XX OS  
XX PN WO200153312-A1.  
XX PD 26-JUL-2001.  
XX PF 26-DEC-2000; 2000WO-US34263.  
XX PR 21-JAN-2000; 2000US-0488725.  
XX PR 25-APR-2000; 2000US-0552317.  
XX PR 09-JUL-2000; 2000US-0598042.  
XX PR 19-JUL-2000; 2000US-0620312.  
XX PR 03-AUG-2000; 2000US-0653450.  
XX PR 14-SEP-2000; 2000US-0662191.  
XX PR 19-OCT-2000; 2000US-0693036.  
XX PR 29-NOV-2000; 2000US-0727344.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Zhao J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX WPI; 2001-442253/47.  
XX P-PSDB; AAM41268.  
XX PT Novel nucleic acids and polypeptides, useful for treating disorders  
XX PT such as central nervous system injuries -  
XX PS Claim 1; SEQ ID NO 4413; 10078pp; English.  
XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
XX CC the encoded polypeptides (AAM38642-AAM42213) with neotropic,  
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide  
XX CC of the invention may be used to treat diseases of the peripheral nervous  
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and  
XX CC localised neuropathies and central nervous system diseases, such as  
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.

XX SQ Sequence 2164 BP; 557 A; 542 C; 560 G; 505 T; 0 other;

Alignment Scores: 4.05e-97 Length: 2164  
Pred. NO.: 1246.00 Matches: 237  
Score: 1246.00  
Percent Similarity: 85.49% Conservative: 40  
Best Local Similarity: 73.15% Mismatches: 45  
Query Match: 49.58% Indels: 2  
DB: 22 Gaps: 2

US-09-960-643-2 (1-476) x AAI60424 (1-2164)

QY 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20  
|||  
Db 144 ATGCCCGGAGAACGGCGAGAGAGCTCTCTGGAAAAAGCAAGCTGAAGACATCAAG 203  
|||  
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
|||  
Db 204 AAGATCTTCGAGTTCAAAGAGACCTCGGAACCGGGGCTTTCCGAAGTGTGTTAGCT 263  
|||  
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59  
|||||  
Db 264 GAAGAGAAGGCAAGCTGGCAAGCTCTTTGCTGTAAGTGATCCCTAAGAGGCGCTGAAG 323  
|||  
QY 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79  
|||||  
Db 324 GCAAGGAAGAACGATAGACAGATGAGATAGCCGCTCTGAGAAGATTAAGCATGAAAT 383  
|||  
QY 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99  
|||||  
Db 384 ATTGTTGCCCTGGAAGACATTTATGAAGCCCAATCACCTGTACTTGGTTCATGCAGCTG 443  
|||  
QY 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119  
|||||  
Db 444 GTGTCCGGTGGAGAGCTGTTTGACCGGATAGTGGAGAAGGGGTTTATACAGAAGAT 503  
|||  
QY 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139  
|||||  
Db 504 GCCAGACTCTGATCCGCCAAGTCTTGGACGCGGTGTACTATCTCCACAGAATGGCATC 563  
|||  
QY 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159  
|||||  
Db 564 GTCCACAGAGACCTCAAGCCGAAAAATCTCTGTACTACATCAAGATGAGGAGTCCAAA 623  
|||  
QY 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178  
|||||  
Db 624 ATAATGATCAGTACTTTGGATTGTCAAAATGGAGGCAAGAGAGATGTGATGCTCACT 683  
|||  
QY 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198  
|||||  
Db 684 GCCTGTGGAACCTCCAGGCTATGTCGCTCTCAAGTCTCTGCCCAAGAACCTTACAGCAA 743  
|||  
QY 199 AlaValAspCysTrpSerIleGlyValIleThrThrIleLeuLeuCysGlyTyrProPro 218  
|||||  
Db 744 GCCGTTGACTGCTGCTCCATCGAGTGATTCCTACATCTTGTCTCGGGCTACCCCTCT 803  
|||  
QY 219 PheTyrGluThrGluSerLysLeuPheGluLysIleLysGlyTyrTyrGluPhe 238  
|||||  
Db 804 TTTATGATGAATGATCTCCAGCTCTTTGAGCAGATCTCTCAAGCGGGAATATAGTTT 863  
|||  
QY 239 GluSerProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258  
|||||  
Db 864 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAGACTTTCATTCGGAACCTGATG 923  
|||



|||||  
Db 995 GAGAGGAGCCCAATAAAGATACACGTGTGAGCAGGAGCGTGGCCACCCATGGATCGCT 1054  
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298  
Db 1055 GGTGACACAGCCCTCAACAAAACATCCACGAGTCCGTCCAGCGCCAGATCCGGAANAAC 1114  
Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318  
Db 1115 TTTCGACAGACCAATGGAGACAGCATTTAATGCCAGCGCGCTGCTGAGACATATGAGA 1174  
Qy 319 LysLeuHisMet 322  
Db 1175 AAACATACACCTC 1186  
RESULT 14  
AAC90432  
ID AAC90432 standard; cDNA; 1694 BP.  
XX  
AC AAC90432;  
XX  
DT 19-MAR-2001 (first entry)  
XX  
DE Murine Dendritic Cell Kinase 3 coding sequence.  
XX  
KW Murine; Dendritic Cell Kinase; MDCK-3; autoimmune disorder; allergy;  
KW wound healing; periodontal disease; inflammatory disease; tumour;  
KW infection; Ca2+/calmodulin-dependent kinase family; ss.  
XX  
OS Mus musculus.  
XX  
PN WO200073468-A1.  
XX  
PD 07-DEC-2000.  
XX  
PF 26-MAY-2000; 2000WO-US14696.  
XX  
PR 28-MAY-1999; 99US-0136781.  
XX  
PA (IMMV) IMMUNEX CORP.  
XX  
PI Bird TA, Virca GD, Martin U, Anderson DM;  
XX  
DR WPI; 2001-061546/07.  
XX  
PT P-PSDB; AAB50055.  
XX  
PT Novel murine and human kinase nucleic acids useful for treating  
XX inflammations, infections, tumors, allergies, autoimmune diseases, and  
XX for stimulating or suppressing immune responses -  
XX  
PS Claim 1; Page 86; 106pp; English.  
XX  
CC The present sequence is the coding sequence for Murine Dendritic Cell  
CC Kinase 3 (MDCK-3). The protein encoded by the present sequence is useful  
CC for treating a variety of disorders listed in the disclosure of the  
CC specification, including autoimmune disorders, allergic reactions,  
CC myeloid or lymphoid cell deficiencies, wound healing and tissue repair  
CC and replacement, burns, incisions and ulcers, periodontal disease,  
CC inflammatory diseases, tumours and bacterial, viral or fungal infection.  
XX MDCK-3 is a member of the Ca2+/calmodulin-dependent kinase family.  
XX  
SQ Sequence 1694 BP; 437 A; 444 C; 464 G; 349 T; 0 other;  
Alignment Scores:  
Pred. No.: 1,26e-96 Length: 1694  
Score: 1238.50 Matches: 258  
Percent Similarity: 68.74% Conservative: 52  
Best Local Similarity: 57.21% Mismatches: 90  
Query Match: 49.28% Indels: 51  
DB: 22 Gaps: 9  
US-09-960-643-2 (1-476) x AAC90432 (1-1694)

Qy 1 MetGlyArgLysGluGluAspAspCysSerTrpLysLysGlnThrThrAsnIleArg 20  
Db 243 ATGCCCGGAGAGAGCGGAGAGCAGCTCTCTCGAANAAGCAAGACAGACATTAAG 302  
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
Db 303 AAGATCTTCAGTTCGAAGAGAGCCCTCGGAACCTGGGCCCTTTCTGAAGTGTGTTAGCC 362  
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle--LysLysSerProAla 59  
Db 363 GAGAGAGAGCTACTGGGAAGCTCTTCGAGTGAAGTGCATCCCGAAGAAGGCGCTGAAG 422  
Qy 60 PheArgAspSerSerLeuGluAsnGlnIleAlaValLeuLysLysIleLysHisGluAsn 79  
Db 423 GGCAAGGAGCAGCAGCATCGAAGAGAGATTGCCGTGCTTAGAAGATTAAAGCATGAAAC 482  
Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99  
Db 483 ATTGTTGCTTGAAGATATTTATGAAGCCCAAAATCACCTCTACCTGGTGCATGCAACTT 542  
Qy 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119  
Db 543 GTGTCTGTGGAGAACTCTTCGATCGGTAGTAGGGAGAGGGGTTTACACAGAGAAGAT 602  
Qy 120 AlaSerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139  
Db 603 GCCAGCACTCTCATCCGCCAGGTCTCGATGCCGTATATCTATCTCCACAGAAATGGGCATT 662  
Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159  
Db 663 GTCCACAGGAGACCTCAAGCCGAGATCTCTTATCTATCTACAGTCAAGACGAGGAGTCCAAA 722  
Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178  
Db 723 ATAATGATCAGTGACTTTGGCTTGTGAAATGGAGGCAAGAGAGATGTGATGCCAG 782  
Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198  
Db 783 GCCTGCGGAGACCCAGGCTATGTTGCTCCGGAAGTTCTCGCCCAAGAACGTCACAGAAA 842  
Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218  
Db 843 GCTGTGACTGCTGCTCCATCGGGTGATCGCTATATATCTTGTCTGTGTGTACCTCTCT 902  
Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238  
Db 903 TTTTATGATGAAATGACTCGAAGCTCTTTGAACAGATCTCTCAAGGCAAGATATAGCTTT 962  
Qy 239 GluSerProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258  
Db 963 GATTCCCTCTACTGGGATGACATCTCCGACTCTGCCAAGACTTCATTCGGAATCTGATG 1022  
Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTyrIleAsp 278  
Db 1023 GAGAAAGACCCAAATAAAGATACACTTGTGAGCAGGACGCTCGACACCCATGGATTGCT 1082  
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298  
Db 1083 GGTGACACAGCCCTTAGCAAAAACATTCACGAATCTCTAGTGCAGCCAGATCCGGAAGAAT 1142  
Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArg 318  
Db 1143 TTTCGAAAGAGCAAAATGGAGCAAGCGTTTAAACGCCAGCGGCTGCTGAGACATATCGG 1202  
Qy 319 LysLeuHisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProPro 338  
Db 1203 AGGCTCCAGCTT-----GGCAGCAGCTGGACAGCTTCAATGCAAGTGTCT 1247  
Qy 339 GluThrGlnAlaSerGluThrArgProSerSerProGluIleThrIle----- 355  
Db 1248 TCAAGCAACCTCAGTTTGGCCAGCAAAAAGATTGTGCGTCTGGCACCTCTCCACGCTCTG 1307  
Qy 356 -----ThrGluAla 358



```

QY 205 ileGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyrGluGluThrGlu 224
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
615 ATCGAGTGATTGCGCTACATCTGCTCGGCTACCCCTCTTTTATGATGAAATGAC 674
QY 225 SerLysLeuPheGluLysGluGlyTyrTyrGluPheGluSerProPheTrpAsp 244
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
675 TCCAAGCTCTTTGACGACATCCTCAAGCGCGAATATGAGTTGACTCTCCCTACTGGAT 734
QY 245 AspileSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLysAspProAsnGlu 264
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
735 GACATCTCGGACTCTGCAAGACTTCATTCCGAACTGATGGAGAGGCCCGCAATAAA 794
QY 265 ArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsnThrAlaLeuHis 284
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
795 AGATACAGTGTGACGACGAGCTCGGCACCCATGGATCGGTGGTGACACAGCCCTCAAC 854
QY 285 ArgAspileTyrProSerValSerLeuGlnIleGlnLysAsnPheAlaLysSerLysTrp 304
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
855 AAAAAATCCACGAGTCCGTCAGCGCCAGATCCGGAATACTTTGCCAAGAGCAATGG 914
QY 305 ArgGlnAlaPheAsnAlaAlaValValHisMetArgLysLeuHisMet 322
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
915 AGACAAGCATTTAATGCCACGCCCTCGTGACACATATGAGAAACTACACCTC 968

```

Search completed: March 15, 2003, 05:11:17  
Job time : 344 secs



QY 293 TCACAGATGAACATTTGTGACCTCGAGAGACATCTATGAGAGACACCCACTACTACC 352  
Db 241 TCAAGCATGAACATTTGTGACCTCGAGAGACATCTATGAGAGACACCCACTACTACC 300  
QY 353 TGGTCATGACAGTGTGTTCTGTTGGGAGCTCTTTGACCGGATCTCGAGCGGGTGCT 412  
Db 301 TGGTCATGACAGTGTGTTCTGTTGGGAGCTCTTTGACCGGATCTCGAGCGGGTGCT 360  
QY 413 ACACAGAGAAGATGACGATCTGGTATCCAGCAGGTCTTTGGCAGTGGAATACCTAC 472  
Db 361 ACACAGAGAAGATGACGATCTGGTATCCAGCAGGTCTTTGGCAGTGGAATACCTAC 420  
QY 473 ATGAGATGACATGCTCCACAGAGACTTAAGCCCGAAGACCTCTTACCTTACCCCTG 532  
Db 421 ATGAGATGACATGCTCCACAGAGACTTAAGCCCGAAGACCTCTTACCTTACCCCTG 480  
QY 533 AAGAGAACTCTAAGATCATGATCACTGACTTTGGTCTGTCCAGATGGAACAGATGSCA 592  
Db 481 AAGAGAACTCTAAGATCATGATCACTGACTTTGGTCTGTCCAGATGGAACAGATGSCA 540  
QY 593 TCATGTCACATGCTGTGGACCCAGGCTACGTGGCTCCAGAGTGTGGCCCGAGAAG 652  
Db 541 TCATGTCACATGCTGTGGACCCAGGCTACGTGGCTCCAGAGTGTGGCCCGAGAAG 600  
QY 653 CCTACAGCAAGGCTGTGGATTGCTGCTTCCATCGGCTCATCACTACATATTGCTCTGTG 712  
Db 601 CCTACAGCAAGGCTGTGGATTGCTGCTTCCATCGGCTCATCACTACATATTGCTCTGTG 660  
QY 713 GATACCCCGCTTCTATGAAGAAGCGAGTCTTAAGCTTTTCGAGAGATCAAGGAGGCT 772  
Db 661 GATACCCCGCTTCTATGAAGAAGCGAGTCTTAAGCTTTTCGAGAGATCAAGGAGGCT 720  
QY 773 ACTATGATTTGAGTCTCCATTCTGGATGACATTTCTGAGTCAGCAAGGACTTTATT 832  
Db 721 ACTATGATTTGAGTCTCCATTCTGGATGACATTTCTGAGTCAGCAAGGACTTTATT 780  
QY 833 GCCACTTGTGTTAGAAGATCCGAAGAGCGGTACACCTGTGAGAGGCGTTGAGTCATC 892  
Db 781 GCCACTTGTGTTAGAAGATCCGAAGAGCGGTACACCTGTGAGAGGCGTTGAGTCATC 840  
QY 893 CCTGGATTGCGGAACACACCGCCCTCCACCGGGACATCTACCANTCAGTCAGCTCCAGA 952  
Db 841 CCTGGATTGCGGAACACACCGCCCTCCACCGGGACATCTACCANTCAGTCAGCTCCAGA 900  
QY 953 TCCAGAGAACTTTGCTTAAGAGCAAGTGGAGCAAGCCTTCAACGACAGCTGTGTGTC 1012  
Db 901 TCCAGAGAACTTTGCTTAAGAGCAAGTGGAGCAAGCCTTCAACGACAGCTGTGTGTC 960  
QY 1013 ACCACATGAGGAAGCTACATGAACCTGCACAGCCCGGCTCCGCCCGAGAGTGGAGA 1072  
Db 961 ACCACATGAGGAAGCTACATGAACCTGCACAGCCCGGCTCCGCCCGAGAGTGGAGA 1020  
QY 1073 ACAGCCGCTTGAACCTCAAGCCTCAGAACTCTAGACCAGCTCCCTGAGATCAGCA 1132  
Db 1021 ACAGCCGCTTGAACCTCAAGCCTCAGAACTCTAGACCAGCTCCCTGAGATCAGCA 1080  
QY 1133 TCACCGAGGACCTGCTCCTGGACACAGTGTAGCACTTCCCTGCGCTGACCCCAATTAACCT 1192  
Db 1081 TCACCGAGGACCTGCTCCTGGACACAGTGTAGCACTTCCCTGCGCTGACCCCAATTAACCT 1140  
QY 1193 GCCAGATGGCGCGCGCCACCTGCGCTGGGAGGTCCTTCAACTGCTGTTCAATG 1252  
Db 1141 GCCAGATGGCGCGCGCCACCTGCGCTGGGAGGTCCTTCAACTGCTGTTCAATG 1200  
QY 1253 GCTCCCTCCACATCAGCAGAGGCTGTGTCCTATGATCAGGSGTCCCTGGCCCGCGGCG 1312  
Db 1201 GCTCCCTCCACATCAGCAGAGGCTGTGTCCTATGATCAGGSGTCCCTGGCCCGCGGCG 1260  
QY 1313 CTTGTGCTGTGCTCAGCTGCTGACATTTGGGAGCAAGGAGTCTCTACTGCT 1372  
Db 1261 CTTGTGCTGTGCTCAGCTGCTGACATTTGGGAGCAAGGAGTCTCTACTGCT 1320

QY 1373 CTGAGCCCACTCTCTCAAAAGGCCAACAAAAACAGAACTTCAAGTCGAGGTATGG 1432  
Db 1321 CTGAGCCCACTCTCTCAAAAGGCCAACAAAAACAGAACTTCAAGTCGAGGTATGG 1380  
QY 1433 TACCAGTTAAAGCCAGTGGCAGTCCCACTCCCGGAGGAGGAGTGGAGTCTGTCTCA 1492  
Db 1381 TACCAGTTAAAGCCAGTGGCAGTCCCACTCCCGGAGGAGGAGTGGAGTCTGTCTCA 1440  
QY 1493 TTATGTGATCTCTGAGGCTGTGCTATGTCACTGCAATTTTCAGGAGACATATTCACT 1552  
Db 1441 TTATGTGATCTCTGAGGCTGTGCTATGTCACTGCAATTTTCAGGAGACATATTCACT 1500  
QY 1553 CTTCTCTCTTCCAAACCTGCTATCCCGCAGAGAGGAGGAGGAGGAGGAGTGGAG 1612  
Db 1501 CTTCTCTCTTCCAAACCTGCTATCCCGCAGAGGAGGAGGAGGAGGAGGAGTGGAG 1560  
QY 1613 CAGGGCTTAGCAGGAGAGTCTTCTGGCCAGAGACACAGCTGTGCTGCGAGGGGAGCC 1672  
Db 1561 CAGGGCTTAGCAGGAGAGTCTTCTGGCCAGAGACACAGCTGTGCTGCGAGGGGAGCC 1620  
QY 1673 CCTCATAGGAGGCCCGCAGGAGGAGGCCCAAGGCTTAGAAGCTTTGTTGAAGCTGTGAGCA 1732  
Db 1621 CCTCATAGGAGGCCCGCAGGAGGAGGCCCAAGGCTTAGAAGCTTTGTTGAAGCTGTGAGCA 1680  
QY 1733 GGAGAGCGGTGCCACACAGCTTCAGGTCTCCCTGACCTGCCCTATGCCCCAC 1790  
Db 1681 GGAGAGCGGTGCCACACAGCTTCAGGTCTCCCTGACCTGCCCTATGCCCCAC 1738

RESULT 2

US-09-935-464-2  
; Sequence 2, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCH.  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300  
; PRIOR FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 1383  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-935-464-2

Query Match 50.6%; Score 1239; DB 9; Length 1383;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1339; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 70 ATGGGTGCGAAGGAAGAGATGACTGCGAGTCTCGAAGAAACAGACACCAACATCCG 129  
Db 1 ATGGGTGCGAAGGAAGAGATGACTGCGAGTCTCGAAGAAACAGACACCAACATCCG 60  
QY 130 AAAACCTTCATTTTATGGAAGTGTGGGATCAGGAGCTTTCTCAGAAGTTTCTCGTG 189  
Db 61 AAAACCTTCATTTTATGGAAGTGTGGGATCAGGAGCTTTCTCAGAAGTTTCTCGTG 120  
QY 190 AAGCAAGACTGACTGGGAGCTCTTGTCTGAGAGTGCATCAAGAGTCACTGCGCTTC 249  
Db 121 AAGCAAGACTGACTGGGAGCTCTTGTCTGAGAGTGCATCAAGAGTCACTGCGCTTC 180  
QY 250 CGGACAGCAGGCTGAGAGATGAGATTGCTGTGTTGAAAAGATCAAGCATGAAAACATT 309  
Db 181 CGGACAGCAGGCTGAGAGATGAGATTGCTGTGTTGAAAAGATCAAGCATGAAAACATT 240  
QY 310 GTGACCTTGAGAGACATCTATGAGAGCACCCCACTACTACCTGGTGCATGCAGCTGTT 369

Db 241 GTGACCTGGAGACATCTATGAGACACACCCACTACTACTGGTCAATGACGTTGTT 300  
QY 370 TCTGGTGGGAGCTCTTTGACCGGATCCCTGGAGCGGGTGTCTACACAGAGAAGATGCC 429  
Db 301 TCTGGTGGGAGCTCTTTGACCGGATCCCTGGAGCGGGTGTCTACACAGAGAAGATGCC 360  
QY 430 AGTCTGGTGTATCCAGAGGCTTGTGCGGAGTGAATACCTACATGAGAATGGCATGCTC 489  
Db 361 AGTCTGGTGTATCCAGAGGCTTGTGCGGAGTGAATACCTACATGAGAATGGCATGCTC 420  
QY 490 CACAGACATTAAGCCCGAAACCTGCTTACCTTACCCCTGAAGAGAACTCTAAGATC 549  
Db 421 CACAGACATTAAGCCCGAAACCTGCTTACCTTACCCCTGAAGAGAACTCTAAGATC 480  
QY 550 ATGATCACTGACTTTGGTCTGTCCAAGATGGAACAGATGGCATGTCCTACCTGCCGT 609  
Db 481 ATGATCACTGACTTTGGTCTGTCCAAGATGGAACAGATGGCATGTCCTACCTGCCGT 540  
QY 610 GGGACCCAGGCTACCTGGCTCCAGAACTGCTGGCCGAAACCCCTACAGAAAGCTGTG 569  
Db 541 GGGACCCAGGCTACCTGGCTCCAGAACTGCTGGCCGAAACCCCTACAGAAAGCTGTG 600  
QY 670 GATTGCTGGTCCATCGGCTCATCACTACATATTGCTGTGGATACCCCGCTTCTAT 729  
Db 601 GATTGCTGGTCCATCGGCTCATCACTACATATTGCTGTGGATACCCCGCTTCTAT 560  
QY 730 GAAGAAACGAGTCTTAAGCTTTTCGAGAGATCAAGAGGGGTACTATGAGTTTGAGTCT 789  
Db 661 GAAGAAACGAGTCTTAAGCTTTTCGAGAGATCAAGAGGGGTACTATGAGTTTGAGTCT 720  
QY 790 CCATTCTGGATCACATTTCTGAGTCAGCCAGGACTTTATTGGCCACTTGTCTTGAGAG 849  
Db 721 CCATTCTGGATCACATTTCTGAGTCAGCCAGGACTTTATTGGCCACTTGTCTTGAGAG 780  
QY 850 GATCCGAACAGCGGTACACCTGTGAGAGAGGCTTGTGAGTCATCCCTGATTCACGGAAC 909  
Db 781 GATCCGAACAGCGGTACACCTGTGAGAGAGGCTTGTGAGTCATCCCTGATTCACGGAAC 840  
QY 910 ACGGCCCTCCACGGGACATCTACCCATCAGTCAGGCTCCAGTCCAGAGAAGCTTTGCT 969  
Db 841 ACAAGCCCTCCACGGGACATCTACCCATCAGTCAGGCTCCAGTCCAGAGAAGCTTTGCT 900  
QY 970 AAGAGAAAGTGGAGCAAGCTTCAACGAGCAGCTGTGGTGCACACATAGGAAGTCA 1029  
Db 901 AAGAGCAAGTGGAGCAAGCTTCAACGAGCAGCTGTGGTGCACACATAGGAAGTCA 960  
QY 1030 CACATGAACCTGCACAGCCCGGCGTCCGCCAGAGGTTGGAGACAGCGCGCTGAAACT 1089  
Db 961 CACATGAACCTGCACAGCCCGGCGTCCGCCAGAGGTTGGAGAAACAGGCGCGCTGAAACT 1020  
QY 1090 CAAGCCTCAAGAACCTCTAGACCCAGCTCCCTGAGATCACCATCACCGAGGACCTGTC 1149  
Db 1021 CAAGCCTCAAGAACCTCTAGACCCAGCTCCCTGAGATCACCATCACCGAGGACCTGTC 1080  
QY 1150 CTGACCCACAGTGTAGCACTCCCTGCGCTGAGCCCAATTAACCTGCGCAGCATGGCGCGG 1209  
Db 1081 CTGACCCACAGTGTAGCACTCCCTGCGCTGAGCCCAATTAACCTGCGCAGCATGGCGCGG 1140  
QY 1210 CCCACTGCCCTGGTGGCAGGTCCTCAACTGCTGGTCAATGGGTCCCTCCACATFACG 1269  
Db 1141 CCCACTGCCCTGGTGGCAGGTCCTCAACTGCTGGTCAATGGGTCCCTCCACATFACG 1200  
QY 1270 AGCAGCTGGTGGCCATGATCAGGGGTCCCTGGCGCGCGGCTGTGGCTGCTGCC 1329  
Db 1201 AGCAGCTGGTGGCCATGATCAGGGGTCCCTGGCGCGCGGCTGTGGCTGCTGCC 1260  
QY 1330 AGCTGCTGAACATTTGGGAGCAAGGAAAGTCTCTACTGCTGTGAGCCCACTCTCTC 1389  
Db 1261 AGCTGCTGAACATTTGGGAGCAAGGAAAGTCTCTACTGCTGTGAGCCCACTCTCTC 1320  
QY 1390 AAAAGGCCCAACAAAAACAG 1410  
|||||

Db 1321 AAAAAGGCCAACAAAAACAG 1341  
RESULT 3  
US-09-935-464-1  
; Sequence 1, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSY  
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300  
; PRIOR FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 157875  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-935-464-1  
Query Match 36.9%; Score 903; DB 9; Length 157875;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1535 CAGGAGACATATTCACCTCTCTCTCCAAACCTGGTGTCTATCCGGCAGAGGGAGG 1594  
Db 147496 CAGGAGACATATTCACCTCTCTCTCCAAACCTGGTGTCTATCCGGCAGAGGGAGG 147555  
QY 1595 AAGCAGACAGCAAGTGGAGCAGGGCTTAGCAGGAGCAGTCTTGGCCAGAGCACCAGCCT 1654  
Db 147556 AAGCAGACAGCAAGTGGAGCAGGGCTTAGCAGGAGCAGTCTTGGCCAGAGCACCAGCCT 147615  
QY 1655 GCTGCCAGCGGGCAGCGCCCTCATAGGAGGCCAGGAGGAGCCCAAGCGGTAGAGCC 1714  
Db 147616 GCTGCCAGCGGGCAGCGCCCTCATAGGAGGCCAGGAGGAGCCCAAGCGGTAGAGCC 147675  
QY 1715 TTGTTGAAGCTGTGAGCAGAGAAAGCGGTGCCACACAGCTTCCAGGTCTCCCTGACCTGC 1774  
Db 147676 TTGTTGAAGCTGTGAGCAGAGAAAGCGGTGCCACACAGCTTCCAGGTCTCCCTGACCTGC 147735  
QY 1775 CTGCTCTATGCCCCACACCCCTACGTCGGTGTCTGTGCAGTGTACGTAGATGCTCTC 1834  
Db 147736 CTGCTCTATGCCCCACACCCCTACGTCGGTGTCTGTGCAGTGTACGTAGATGCTCTC 147795  
QY 1835 GCCTGGGTCTGTGCTGTTGTCGTGAAAGCTTAATGGGCTGGCCAGGCTGTGTCACCTT 1894  
Db 147796 GCCTGGGTCTGTGCTGTTGTCGTGAAAGCTTAATGGGCTGGCCAGGCTGTGTCACCTT 147855  
QY 1895 CTCCAAGCAAAAGCCATATGGAGCATCTACCAGACTCCCACTCTGCACACACTCACTCCC 1954  
Db 147856 CTCCAAGCAAAAGCCATATGGAGCATCTACCAGACTCCCACTCTGCACACACTCACTCCC 147915  
QY 1955 ACCTCTCAAGCCTTCAAGCTTGGCCAGATTTGGGCTCAATTAATGTCGTCGCTGCCAT 2014  
Db 147916 ACCTCTCAAGCCTTCAAGCTTGGCCAGATTTGGGCTCAATTAATGTCGTCGCTGCCAT 147975  
QY 2015 CTGCATGAATCAGAGGAGCTCCCACTGGTGTGCTGTGAGCTCTTCAAGTTCTTAAT 2074  
Db 147976 CTGCATGAATCAGAGGAGCTCCCACTGGTGTGCTGTGAGCTCTTCAAGTTCTTAAT 148035  
QY 2075 CTTTAACTCCAGGATTAGCTCCCAAGTCGCTGAGACCCAGCCAGCAGCAGCTTCTGGCCCT 2134  
Db 148036 CTTTAACTCCAGGATTAGCTCCCAAGTCGCTGAGACCCAGCCAGCAGCAGCTTCTGGCCCT 148095  
QY 2135 TCTCCCTGCCCTCAATCTAAAAGAGAGTGCCACACCCCTCCAAAGTGGAAATAGAAAGATTC 2194  
Db 148096 TCTCCCTGCCCTCAATCTAAAAGAGAGTGCCACACCCCTCCAAAGTGGAAATAGAAAGATTC 148155  
|||||

QY 2195 ATGAGTAAGGCTGCAAGGAATCTTATCTGCGCACATGCTCCTCGTGACACACCCAA 2254  
|||||  
Db 148156 ATGAGTAAGGCTGCAAGGAATCTTATCTGCGCACATGCTCCTCGTGACACACCCAA 148215  
|||||  
QY 2255 TGGAGTTAACCTTGGAAAGTTGACTATTTTAATGCTGCCAGAGTTCTTAATCCTGCCCTCT 2314  
|||||  
Db 148216 TGGAGTTAACCTTGGAAAGTTGACTATTTTAATGCTGCCAGAGTTCTTAATCCTGCCCTCT 148275  
|||||  
QY 2315 GTTCCCTTTTCTCTCTGAAAGTCCAGCACACCATTTCTTGCTCTCCCGAGTTCCCTCG 2374  
|||||  
Db 148276 GTTCCCTTTTCTCTCTGAAAGTCCAGCACACCATTTCTTGCTCTCCCGAGTTCCCTCG 148335  
|||||  
QY 2375 CCTCCACCCCTCCAGCTTCATGCTCAGTGTGTGCTTAATAAATGGACATATTTTCT 2434  
|||||  
Db 148336 CCTCCACCCCTCCAGCTTCATGCTCAGTGTGTGCTTAATAAATGGACATATTTTCT 148395  
|||||  
QY 2435 CTA 2437  
|||  
Db 148396 CTA 148398

RESULT 4

US-09-935-464-46  
; Sequence 46, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA  
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; PRIOR FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 46  
; LENGTH: 480  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: n=a or g or c or t/u  
US-09-935-464-46

Query Match 18.9%; Score 462; DB 9; Length 480;  
Best Local Similarity 100.0%; Pred. No. 4.7e-218;  
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 59 CTCCTGAGGCATGGGTCCAAAGGAAGATGACTGCAGTTCCTCGGAAGAAACAGACCA 118  
|||||  
Db 19 CTCCTGAGGCATGGGTCCAAAGGAAGATGACTGCAGTTCCTCGGAAGAAACAGACCA 78  
|||||  
QY 119 CCAACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTCTCAGAAG 178  
|||||  
Db 79 CCAACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTCTCAGAAG 138  
|||||  
QY 179 TTTTCTGCTGAGCAAGACACTGCTGGGAAGCTTCTGCTGAAGTGCATCAAGAAGT 238  
|||||  
Db 139 TTTTCTGCTGAGCAAGACACTGCTGGGAAGCTTCTGCTGAAGTGCATCAAGAAGT 198  
|||||  
QY 239 CACCTCCCTTCGGGACACAGCCTGGGAATGAGATGAGTGTGTGAAAAGATCAAGC 298  
|||||  
Db 199 CACCTCCCTTCGGGACACAGCCTGGGAATGAGATGAGTGTGTGAAAAGATCAAGC 258  
|||||  
QY 299 ATGAAACATTTGACCCCTGGAGGACATCTATGAGAGCACACCCACTACTACCTGGTCA 358  
|||||  
Db 259 ATGAAACATTTGACCCCTGGAGGACATCTATGAGAGCACACCCACTACTACCTGGTCA 318  
|||||  
QY 359 TGCAGCTTGTCTGCTGGGAGCTCTTTTGACCGGATCTCGGAGCGGGGTGCTACACAG 418  
|||||

Db 319 TGCAGCTTGTCTGCTGGGAGCTCTTTGACCGGATCCTCGAGCGGGGTGCTACACAG 378  
|||||  
QY 419 AGAAGGATGCCAGTCTGATCCAGCAGCTCTTGTGCGCAGTGAATATCCTACATGAGA 478  
|||||  
Db 379 AGAAGGATGCCAGTCTGATCCAGCAGCTCTTGTGCGCAGTGAATATCCTACATGAGA 438  
|||||  
QY 479 ATGCATCGTCCACAGAGACTTAAAGCCCGAAACCTGCTTT 520  
|||||  
Db 439 ATGCATCGTCCACAGAGACTTAAAGCCCGAAACCTGCTTT 480  
|||||

RESULT 5

US-09-935-464-7  
; Sequence 7, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA  
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; PRIOR FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300  
; PRIOR FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 7  
; LENGTH: 501  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-935-464-7

Query Match 18.4%; Score 450; DB 9; Length 501;  
Best Local Similarity 99.8%; Pred. No. 4e-212;  
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1525 CTGCAATTTTCAGGAGACATATTCAACTCCTCTCTTCCAAACCTGGTGTCTATCCGG 1584  
|||||  
Db 1 CTGCAATTTTCAGGAGACATATTCAACTCCTCTCTTCCAAACCTGGTGTCTATCCGG 60  
|||||  
QY 1585 CAGAGGGAGGAAGCAGAGCAAGTGGAGAGGGCTTAGCAGGAGCAGCTTTCTGGCCAGAA 1644  
|||||  
Db 61 CAGAGGGAGGAAGCAGAGCAAGTGGAGAGGGCTTAGCAGGAGCAGCTTTCTGGCCAGAA 120  
|||||  
QY 1645 GCACCAGCTGCTGCCAGCGGGCAGCCCTCATAGAGGCCAGGAGGAGGCCCAAGG 1704  
|||||  
Db 121 GCACCAGCTGCTGCCAGCGGGCAGCCCTCATAGAGGCCAGGAGGAGGCCCAAGG 180  
|||||  
QY 1705 CGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACCAGCTTCCAGGTCTC 1764  
|||||  
Db 181 CGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACCAGCTTCCAGGTCTC 240  
|||||  
QY 1765 CCTGACCTGCTGCTCTATGCCCCACACCTACGTGCGGTGGCTCTGTGAGTGTAGCTA 1824  
|||||  
Db 241 CCTGACCTGCTGCTCTATGCCCCACACCTACGTGCGGTGGCTCTGTGAGTGTAGCTA 300  
|||||  
QY 1825 GATAGCTCTCCCGTGGGTCTGTGCTTTGTGCGTGAAGCTTAATGGGCTGGCCAGGCT 1884  
|||||  
Db 301 GATAGCTCTCCCGTGGGTCTGTGCTTTGTGCGTGAAGCTTAATGGGCTGGCCAGGCT 360  
|||||  
QY 1885 GTGTACCTTCTCAAGCAAGCCATATGGAGCATCTACCCAGACTCCCCTACTCTGACAC 1944  
|||||  
Db 361 GTGTACCTTCTCAAGCAAGCCATATGGAGCATCTACCCAGACTCCCCTACTCTGACAC 420  
|||||  
QY 1945 ACTCACTCCCACCTCTCAAGCCTCAACCTCTTGCCAGATTTGGGCTCAATTAATGTCTGT 2004  
|||||  
Db 421 ACTCACTCCCACCTCTCAAGCCTCAACCTCTTGCCAGATTTGGGCTCAATTAATGTCTGT 480  
|||||  
QY 2005 GCCTGCCCATCTGCATGAATG 2025  
|||||

```
Db 481 GCCTGCCATCTGCATGAATG 501
RESULT 6
US-09-935-464-50
; Sequence 50, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-50
Query Match 13.0%; Score 319; DB 9; Length 319;
Best Local Similarity 100.0%; Pred. No. 2.1e-147; Indels 0; Gaps 0;
Matches 319; Conservative 0; Mismatches 0;
QY 41 CATCCTCAGAACTTCAACTCTGGAGCAATGGTGCAGAAAGAAAGATGACTGCAGTT 100
Db 1 CATCCTCAGAACTTCAACTCTGGAGCAATGGTGCAGAAAGAAAGATGACTGCAGTT 60
QY 101 CCTGGAAGAAACAGACACCAACATCCGGAACCTTCATTTTATGGAAGTCTGGGAT 160
Db 61 CCTGGAAGAAACAGACACCAACATCCGGAACCTTCATTTTATGGAAGTCTGGGAT 120
QY 161 CAGAGCTTTCTCAGAAAGTTTCTGTTGGAAGCAAGACTGACTGGGAGCTTTTGCTC 220
Db 121 CAGGAGCTTTCTCAGAAAGTTTCTGTTGGAAGCAAGACTGACTGGGAGCTTTTGCTC 180
QY 221 TGAAGTGCATCAAGAAGTCACTGCCCTTCGGGACAGCAGCCCTGGAGATGAGATTGCTG 280
Db 181 TGAAGTGCATCAAGAAGTCACTGCCCTTCGGGACAGCAGCCCTGGAGATGAGATTGCTG 240
QY 281 TGTGAAAAAGATCAAGCATGAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCA 340
Db 241 TGTGAAAAAGATCAAGCATGAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCA 300
QY 341 CCCACTACTACCTGGTCAAT 359
Db 301 CCCACTACTACCTGGTCAAT 319
RESULT 7
US-09-764-868-189
; Sequence 189, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P732
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 189
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: SITE
; LOCATION: (59)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-189
Query Match 11.2%; Score 274; DB 9; Length 476;
Best Local Similarity 100.0%; Pred. No. 3.7e-125; Indels 0; Gaps 0;
Matches 274; Conservative 0; Mismatches 0;
QY 1 TGGAGTGGAGCTCAAGCAGGATTTCCCGAGTCCCTGGGATCCTCAGAAGCTTCAACT 60
Db 145 TGGAGTGGAGCTCAAGCAGGATTTCCCGAGTCCCTGGGATCCTCAGAAGCTTCAACT 204
QY 61 CTGGAGCAATGGTGCAGAAAGAAAGATGACTGCAGTTCCTGGAAGAAACAGACCACC 120
Db 205 CTGGAGCAATGGTGCAGAAAGAAAGATGACTGCAGTTCCTGGAAGAAACAGACCACC 264
QY 121 AACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTT 180
Db 265 AACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTT 324
QY 181 TTCCTGGTGAAGCAAGACTGACTGGGAGCTTTTGTCTCTGAAGTGCATCAAGAACTCA 240
Db 325 TTCCTGGTGAAGCAAGACTGACTGGGAGCTTTTGTCTCTGAAGTGCATCAAGAACTCA 384
QY 241 CCTGCTTCCGGGACAGCAGCTGGAGATGAGA 274
Db 385 CCTGCTTCCGGGACAGCAGCTGGAGATGAGA 418
RESULT 8
US-09-935-464-47
; Sequence 47, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSY
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 47
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: n=a or g or c or t/u
US-09-935-464-47
Query Match 8.7%; Score 212; DB 9; Length 467;
Best Local Similarity 99.4%; Pred. No. 1.6e-94; Indels 0; Gaps 0;
Matches 312; Conservative 0; Mismatches 2;
QY 41 CATCCTCAGAAGCTTCAACTCTGGAGCAATGGTGCAGAAAGAAAGATGACTGCAGTT 100
Db 1 CATCCTCAGAAGCTTCAACTCTGGAGCAATGGTGCAGAAAGAAAGATGACTGCAGTT 60
QY 101 CTGGAAGAAACAGACACCAACATCCGGAACCTTCATTTTATGGAAGTCTGGGAT 160
Db 61 CTGGAAGAAACAGACACCAACATCCGGAACCTTCATTTTATGGAAGTCTGGGAT 120
QY 161 CAGGAGCTTTCTCAGAAGTTTCTGTTGGAAGCAAGACTGACTGGGAGCTTTTGCTC 220
Db 121 CAGGAGCTTTCTCANAAGTTTCTGTTGGAAGCAAGACTGACTGGGAGCTTTTGCTC 180
QY 221 TGAAGTGCATCAAGAACTCACCTTCCGGGACAGCAGCTGGAGATGAGATTGCTG 280
```

Db 181 TGAAGTCATCAAGAAAGTCACCTGCCTTCGGGACAGCAGCCTGNAGAATGAGATTGCTG 240  
Qy 281 TGTGTAAGAAAGATCAAGCATGAAACATGTTGACCCCTGGAGGACATCTATGAGAGCACCA 340  
Db 241 TGTGTAAGAAAGATCAAGCATGAAACATGTTGACCCCTGGAGGACATCTATGAGAGCACCA 300  
Qy 341 CCCACTACTACCTG 354  
Db 301 CCCACTACTACCTG 314

RESULT 9  
US-09-935-464-48  
; Sequence 48, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; PRIOR FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 48  
; LENGTH: 470  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: n=a or g or c or t/u  
US-09-935-464-48

Query Match 8.5%; Score 209; DB 9; Length 470;  
Best Local Similarity 100.0%; Pred. No. 4.8e-93;  
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 183 CTGTGTGAAGCAAGACTGACTGGGAGCTCTTGTCTGGAAGTCATCAAGAAAGTCACCC 242  
Db 146 CTGTGTGAAGCAAGACTGACTGGGAGCTCTTGTCTGGAAGTCATCAAGAAAGTCACCC 205  
Qy 243 TGCCTTCCGGGACAGCAGCCTGGAGATGAGATTGCTGTGTTGAAAAAGATCAAGCATGA 302  
Db 206 TGCCTTCCGGGACAGCAGCCTGGAGATGAGATTGCTGTGTTGAAAAAGATCAAGCATGA 265  
Qy 303 AACATTGTGACCTGGAGACATCTATGAGAGCACCACCTACTACTACCTGGTGCATGCA 362  
Db 266 AACATTGTGACCTGGAGACATCTATGAGAGCACCACCTACTACTACTACCTGGTGCATGCA 325  
Qy 363 GCTGTTTCTGTGGGGAGCTCTTTGACC 391  
Db 326 GCTGTTTCTGTGGGGAGCTCTTTGACC 354

RESULT 10  
US-09-935-464-49  
; Sequence 49, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; PRIOR FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300

; PRIOR FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 49  
; LENGTH: 356  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: n=a or g or c or t/u  
US-09-935-464-49

Query Match 7.6%; Score 185; DB 9; Length 356;  
Best Local Similarity 100.0%; Pred. No. 3.4e-81;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 183 CTGTGTGAAGCAAGACTGACTGGGAGCTCTTGTCTGGAAGTCATCAAGAAAGTCACCC 242  
Db 146 CTGTGTGAAGCAAGACTGACTGGGAGCTCTTGTCTGGAAGTCATCAAGAAAGTCACCC 205  
Qy 243 TGCCTTCCGGGACAGCAGCCTGGAGATGAGATTGCTGTGTTGAAAAAGATCAAGCATGA 302  
Db 206 TGCCTTCCGGGACAGCAGCCTGGAGATGAGATTGCTGTGTTGAAAAAGATCAAGCATGA 265  
Qy 303 AACATTGTGACCTGGAGACATCTATGAGAGCACCACCTACTACTACTACTGGTGCATGCA 362  
Db 266 AACATTGTGACCTGGAGACATCTATGAGAGCACCACCTACTACTACTACTGGTGCATGCA 325  
Qy 363 GCTTG 367  
Db 326 GCTTG 330

RESULT 11  
US-09-731-872-177  
; Sequence 177, Application US/09731872  
; Patent No. US20020102604A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bouqueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78.US3.REG  
; CURRENT APPLICATION NUMBER: US/09/731,872  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 177  
; LENGTH: 1361  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: CDS  
; LOCATION: 65..1024  
US-09-731-872-177

Query Match 6.1%; Score 149; DB 10; Length 1361;  
Best Local Similarity 100.0%; Pred. No. 2.3e-63;  
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2299 TTCTAATCCCTGCCTCTCTTCCCTTTTCTCCCTTAAAGTCCAGACACACCATTTCTGTCC 2358  
Db 1177 TTCTAATCCCTGCCTCTCTTCCCTTTTCTCCCTTAAAGTCCAGACACACCATTTCTGTCC 1236  
Qy 2359 TTCCCCAGTTTCTCGCCCTCCACCCCTCAGCTTCATGCTCAGTGTGTGCTTAATAAA 2418  
Db 1237 TTCCCCAGTTTCTCGCCCTCCACCCCTCAGCTTCATGCTCAGTGTGTGCTTAATAAA 1296  
Qy 2419 ATGGACATATTTTCTCTAAAAA 2447

Db 1297 ATGCACATATTTCTCTCTAAAAA 1325  
|||||

RESULT 12

US-09-935-464-6  
; Sequence 6, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA  
; FILE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300  
; PRIOR FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 386  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: n = a or g or c or t/u  
US-09-935-464-6

Query Match 3.3%; Score 80; DB 9; Length 386;  
Best Local Similarity 100.0%; Pred. No. 2.6e-29; Indels 0; Gaps 0;  
Matches 80; Conservative 0; Mismatches 0;

QY 1523 CACTGCAATTTTCAGGAGACATATTCAACTCTCTCTCTTCCAAACCTGGTGTCTATCC 1582  
|||||

Db 3 CACTGCAATTTTCAGGAGACATATTCAACTCTCTCTCTTCCAAACCTGGTGTCTATCC 62

QY 1583 GGCAGAGGGAGGAGGCAGA 1602  
|||||

Db 63 GGCAGAGGGAGGAGGCAGA 82

RESULT 13

US-09-796-692-7777/C  
; Sequence 7777, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; FILE OF INVENTION: HEMATOLOGICAL MALIGNANCIES  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09/796,692  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: 60/218,950

; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 9597  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7777  
; LENGTH: 474  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (10)  
; OTHER INFORMATION: n=A,T,C or G  
; NAME/KEY: unsure  
; LOCATION: (19)  
; OTHER INFORMATION: n=A,T,C or G  
; NAME/KEY: unsure  
; LOCATION: (20)  
; OTHER INFORMATION: n=A,T,C or G  
US-09-796-692-7777

Query Match 1.3%; Score 32; DB 9; Length 474;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05; Indels 0; Gaps 0;  
Matches 32; Conservative 0; Mismatches 0;

QY 652 CCTACAGCAAGGCTGTGGATTGCTGGTCCAT 683  
|||||

Db 229 CCTACAGCAAGGCTGTGGATTGCTGGTCCAT 198

RESULT 14

US-09-817-181-1  
; Sequence 1, Application US/09817181  
; Patent No. US20020142427A1  
; GENERAL INFORMATION:  
; APPLICANT: MERRILOV, Gennady et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001189  
; CURRENT APPLICATION NUMBER: US/09/817,181  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1372  
; TYPE: DNA  
; ORGANISM: Human  
US-09-817-181-1

Query Match 1.3%; Score 32; DB 10; Length 1372;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05; Indels 0; Gaps 0;  
Matches 32; Conservative 0; Mismatches 0;

QY 652 CCTACAGCAAGGCTGTGGATTGCTGGTCCAT 683  
|||||

Db 582 CCTACAGCAAGGCTGTGGATTGCTGGTCCAT 613

RESULT 15

US-09-935-464-45  
; Sequence 45, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSY  
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA

```

; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 45
; LENGTH: 27
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA probe
US-09-935-464-45

```

```

Query Match      1.1%; Score 27; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 293 TCAAGCATGAAACATTGTGACCTGG 319
Db 1 TCAAGCATGAAACATTGTGACCTGG 27

```

```

Search completed: March 15, 2003, 05:09:04
Job time : 401 secs

```

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:03:36 ; Search time 38 Seconds  
(without alignments)  
1669.139 Million cell updates/sec

Title: US-09-960-643-2

Perfect score: 2513

Sequence: 1 MGRKEDDCSSWKQTNTNR.....VKASGSHCRAGTGVCLIM 476

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A\_Geneseq\_101002:\*
- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
  - 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
  - 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
  - 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
  - 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
  - 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
  - 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
  - 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
  - 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
  - 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
  - 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
  - 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
  - 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
  - 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
  - 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
  - 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
  - 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
  - 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
  - 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
  - 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
  - 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
  - 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
  - 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2513	100.0	476	23	Human calmodulin k
2	2513	100.0	497	22	Human polypeptide
3	2499.5	99.5	567	22	Human polypeptide
4	1382	55.0	298	22	Human diagnostic a
5	1246	49.6	355	22	Human kinase (PKIN
6	1246	49.6	355	22	Human polypeptide
7	1246	49.6	355	23	Human Cam kinase I
8	1246	49.6	357	22	Human kinase (PKIN
9	1246	49.6	385	22	Human polypeptide
10	1246	49.6	385	22	Amino acid sequenc

11	1241	49.4	357	22	AAU03508	Human protein kina
12	1233	49.1	355	22	AAU03505	Murine Dendritic C
13	1231.5	49.0	356	22	AAU03505	Amino acid sequenc
14	1145	45.6	389	21	AAU03505	Amino acid sequenc
15	1070	42.6	358	19	AAU03505	Calmodulin-depende
16	1039	41.3	343	22	AAU03511	Human protein kina
17	1036	41.2	503	22	ABG05970	Novel human diagno
18	990	39.4	343	20	AAU03505	Disease associated
19	957.5	38.1	405	22	ABU08436	Drosophila melanog
20	944.5	37.6	280	22	AAU01034	Human death domain
21	747.5	29.7	501	22	AAU01783	Human kinase (PKIN
22	747.5	29.7	501	22	AAU01783	Amino acid sequenc
23	747.5	29.7	501	22	AAU01783	Human protein kina
24	746.5	29.7	501	22	AAU01783	Human polypeptide
25	745.5	29.7	470	22	AAU03404	Human polypeptide,
26	745.5	29.7	470	22	AAU03404	Human polypeptide,
27	745.5	29.7	470	22	AAU03404	Human protein phos
28	745	29.6	470	22	AAU03404	Human polypeptide
29	733.5	29.2	501	22	AAU03404	Human polypeptide,
30	726.5	28.9	205	22	AAU03404	Human protein sequ
31	691.5	27.5	502	21	AAU03404	A rat calcium/calim
32	650.5	25.9	424	18	AAU03404	PSKH-1 serine kina
33	643.5	25.6	414	22	AAU03404	Human protein SEQ
34	643.5	25.6	482	22	AAU03404	Human protein SEQ
35	639.5	25.4	449	22	AAU03404	Novel human diagno
36	625.5	24.9	766	23	AAU03404	Human PKIN-14 prot
37	625.5	24.9	766	23	AAU03404	Human serine/threo
38	625	24.9	264	20	AAU03404	Human protein kina
39	615.5	24.5	641	21	AAU03404	Amino acid sequenc
40	611.5	24.3	600	22	AAU03404	Human protein kina
41	611.5	24.3	600	23	AAU03404	Human kinase (PKIN
42	603	24.0	474	22	AAU03404	Novel human diagno
43	599.5	23.9	493	22	AAU03404	Drosophila melanog
44	599.5	23.9	493	22	AAU03404	Drosophila melanog
45	599.5	23.9	493	22	AAU03404	Drosophila melanog

ALIGNMENTS

RESULT 1

AAE22764

ID AAE22764 standard; Protein: 476 AA.

XX AAE22764;

XX AAE22764;

DT 09-AUG-2002 (first entry)

XX Human calmodulin kinase, CAMK-XI.

DE Human: cytostatic; antisense gene therapy; screening; protein kinase;

KW cancer; liver; colon; tumour; inflammation; arthritis synovium; CAMK-XI;

KW enzyme; calmodulin kinase.

OS Homo sapiens.

XX WO200224947-A2.

XX 28-MAR-2002.

PF 20-SEP-2001; 2001WO-IB02237.

PR 20-SEP-2000; 2000US-233999P.

PR 02-OCT-2000; 2000US-237419P.

PR 02-OCT-2000; 2000US-237423P.

PR 04-OCT-2000; 2000US-238558P.

PR 10-MAY-2001; 2001US-290555P.

XX (KINE-) KINETEK PHARM INC.

PA (UYBR-) UNIV BRITISH COLUMBIA.

XX Yoganathan T, Delaney AD;

XX

DR WPI: 2002-394145/42.  
DR N-PSDB; AAD36140.  
XX  
PT Diagnosing cancer, comprises determining the upregulation of expression  
PT of a nucleic acid sequence encoding a protein kinase or upregulation of  
PT expression of the protein kinase, in the cancer -  
XX  
XX  
PS Claim 1; Page 64-65; 87pp; English.  
XX  
CC The invention relates to a method for screening biologically active agent  
CC that modulates cancer associated protein kinase function. The invention  
CC also relates to a method for diagnosing cancer comprising determining the  
CC upregulation of expression of a nucleic acid sequence encoding a protein  
CC kinase. The method is useful for diagnosing cancer. A protein kinase is  
CC useful for screening biological agents that modulate cancer associated  
CC protein kinase function. Downregulating the activity of protein kinase is  
CC useful for inhibiting the growth of a cancer cell, e.g. liver or colon  
CC cancer. A nucleic acid encoding protein kinase is useful to screen biopsy  
CC derived tumours and inflammatory samples such as arthritic synovium, for  
CC amplified DNA in the cell or increased expression of corresponding mRNA  
CC or protein and is also useful to detect differences in expression levels  
CC such as molecular weight, amino acid and nucleotide sequences between the  
CC two cells. The present sequence is human calmodulin kinase, CAMK-XI.  
XX  
XX Sequence 476 AA;  
XX  
Query Match 100.0%; Score 2513; DB 23; Length 476;  
Best Local Similarity 100.0%; Pred. No. 4.8e-220;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGRKEEDCSWKKQTNNIRKTFIFMEVLGSGAFSEFVLVKQLTKLFALKCIKSPAF 60  
DB 1 MGRKEEDCSWKKQTNNIRKTFIFMEVLGSGAFSEFVLVKQLTKLFALKCIKSPAF 60  
QY 61 RDSSLENEIAVLKIKHENIVTLEDIYESTTHYLVLMQVSGGELFDRILRGVYTEKDA 120  
DB 61 RDSSLENEIAVLKIKHENIVTLEDIYESTTHYLVLMQVSGGELFDRILRGVYTEKDA 120  
QY 121 SLVIOQVLSAVKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMEQNGIMSTAC 180  
DB 121 SLVIOQVLSAVKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMEQNGIMSTAC 180  
QY 181 GTPGVYAPEVLAQPKYSKAVDCWSIGVITVILLCGYPPFYEETESKLFKIKEGYEFES 240  
DB 181 GTPGVYAPEVLAQPKYSKAVDCWSIGVITVILLCGYPPFYEETESKLFKIKEGYEFES 240  
QY 241 PFWDIDISEAKDFCHLEKDPNERYTCEKALSHPWIDGNTALHRLDIYPSVSLIQKNFA 300  
DB 241 PFWDIDISEAKDFCHLEKDPNERYTCEKALSHPWIDGNTALHRLDIYPSVSLIQKNFA 300  
QY 301 KSKWRQAFNAAVVHHMRKLMNLHSPGVPRPEVENRPPETQASSETSRPSPETITEAPV 360  
DB 301 KSKWRQAFNAAVVHHMRKLMNLHSPGVPRPEVENRPPETQASSETSRPSPETITEAPV 360  
QY 361 LDHVSVALPALTQLPCQHGRRPTAPGGRSLNCLVNGSLHSSIVPMHQGSLAAGPCGCCS 420  
DB 361 LDHVSVALPALTQLPCQHGRRPTAPGGRSLNCLVNGSLHSSIVPMHQGSLAAGPCGCCS 420  
QY 421 SCLNIGSKGSSYCSPTLLKANKKONFKSEYVMPVKASGSHCRAGQTGVCLIM 476  
DB 421 SCLNIGSKGSSYCSPTLLKANKKONFKSEYVMPVKASGSHCRAGQTGVCLIM 476  
RESULT 2  
AAM41547  
ID AAM41547 standard; Protein: 497 AA.  
XX  
AC AAM41547;  
XX  
XX 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 6478.  
XX

Human: nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
peripheral nervous system; neuropathy; central nervous system; CNS;  
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
leukaemia.  
XX Homo sapiens.  
XX  
XX WO200153312-A1.  
XX  
XX 26-JUL-2001.  
XX  
XX 26-DEC-2000; 2000WO-US34263.  
XX  
XX 21-JAN-2000; 2000US-0488725.  
XX 25-APR-2000; 2000US-0552317.  
XX 09-JUL-2000; 2000US-0598042.  
XX 19-JUL-2000; 2000US-0620312.  
XX 03-AUG-2000; 2000US-0653450.  
XX 14-SEP-2000; 2000US-0662191.  
XX 19-OCT-2000; 2000US-0693036.  
XX 29-NOV-2000; 2000US-0727344.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI: 2001-442253/47.  
DR N-PSDB; AAI60703.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
PT  
XX Example 2; SEQ ID NO 6478; 10078pp; English.  
XX  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity; chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity; cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemia and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
XX Sequence 497 AA;  
XX  
Query Match 100.0%; Score 2513; DB 22; Length 497;  
Best Local Similarity 100.0%; Pred. No. 5.1e-220;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGRKEEDCSWKKQTNNIRKTFIFMEVLGSGAFSEFVLVKQLTKLFALKCIKSPAF 60  
DB 22 MGRKEEDCSWKKQTNNIRKTFIFMEVLGSGAFSEFVLVKQLTKLFALKCIKSPAF 81  
QY 61 RDSSLENEIAVLKIKHENIVTLEDIYESTTHYLVLMQVSGGELFDRILRGVYTEKDA 120  
DB 82 RDSSLENEIAVLKIKHENIVTLEDIYESTTHYLVLMQVSGGELFDRILRGVYTEKDA 141  
QY 121 SLVIOQVLSAVKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMEQNGIMSTAC 180  
DB 142 SLVIOQVLSAVKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMEQNGIMSTAC 201

QY 181 GTGYYVAPEVLAQKPYSKAVDCWSIGVITYILLGYPPEYEETESKLFKIKEGYEFPS 240  
|||||  
Db 202 GTGYYVAPEVLAQKPYSKAVDCWSIGVITYILLGYPPEYEETESKLFKIKEGYEFPS 261  
|||||  
QY 241 PFWDIDISEAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRIYPSVSIQIKNFA 300  
|||||  
Db 262 PFWDIDISEAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRIYPSVSIQIKNFA 321  
|||||  
QY 301 KSKWRQAFNAAAVVHHMKLHNLHSPGVRPEVENRPPETQASRSPSPETITTEAPV 360  
|||||  
Db 322 KSKWRQAFNAAAVVHHMKLHNLHSPGVRPEVENRPPETQASRSPSPETITTEAPV 381  
|||||  
QY 361 LDHSHVALPALTPCQHGRRPTAPGGRSLNCLVNGSLHISSSLVPMHQGSLAAGPCGCCS 420  
|||||  
Db 382 LDHSHVALPALTPCQHGRRPTAPGGRSLNCLVNGSLHISSSLVPMHQGSLAAGPCGCCS 441  
|||||  
QY 421 SCLNIGSKGKSSYCEPTLLKANKKQNFSEVMVPVKASGSSHCRAGQTGVCLIM 476  
|||||  
Db 442 SCLNIGSKGKSSYCEPTLLKANKKQNFSEVMVPVKASGSSHCRAGQTGVCLIM 497  
  
RESULT 3  
AAM39761  
ID AAM39761 standard; Protein; 567 AA.  
AC  
XX AAM39761;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 2906.  
XX  
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WC200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US34263.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI; 2001-442253/47.  
XX N-PSDB; AAI58917.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
XX Example 4; SEQ ID NO 2906; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 567 AA:  
  
Query Match 99.5%; Score 2499.5; DB 22; Length 567;  
Best Local Similarity 98.6%; Pred. No. 1.le-218; Indels 7; Gaps 1;  
Matches 476; Conservative 0; Mismatches 0;  
  
QY 1 MGRKEEDCSSWKQTNIIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKSPAF 60  
|||||  
Db 85 MGRKEEDCSSWKQTNIIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKSPAF 144  
|||||  
QY 61 RDSLENETAVLKIKHENIVTLEDIYESTTHYLVLMOLVSGGELEFDRILRGVYTEKDA 120  
|||||  
Db 145 RDSLENETAVLKIKHENIVTLEDIYESTTHYLVLMOLVSGGELEFDRILRGVYTEKDA 204  
|||||  
QY 121 SLVIQOVL SAVKYLHENGIVHRDLK-----PENLLYLTPEENSKIMITDFGLSKMEQN 173  
|||||  
Db 205 SLVIQOVL SAVKYLHENGIVHRDLKVSRRSEWPENLLYLTPEENSKIMITDFGLSKMEQN 264  
|||||  
QY 174 GIMSTACGTPGYVAPVLAQKPYSKAVDCWSIGVITYILLGYPPEYEETESKLFKIKE 233  
|||||  
Db 265 GIMSTACGTPGYVAPVLAQKPYSKAVDCWSIGVITYILLGYPPEYEETESKLFKIKE 324  
|||||  
QY 234 GYEFESPFWDIDISEAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRIYPSVSL 293  
|||||  
Db 325 GYEFESPFWDIDISEAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRIYPSVSL 384  
|||||  
QY 294 QIQKNEAKSKWRQAFNAAAVVHHMKLHNLHSPGVRPEVENRPPETQASRSPSPPEI 353  
|||||  
Db 385 QIQKNEAKSKWRQAFNAAAVVHHMKLHNLHSPGVRPEVENRPPETQASRSPSPPEI 444  
|||||  
QY 354 TITEAPVLDHSHVALPALTPCQHGRRPTAPGGRSLNCLVNGSLHISSSLVPMHQGSLAA 413  
|||||  
Db 445 TITEAPVLDHSHVALPALTPCQHGRRPTAPGGRSLNCLVNGSLHISSSLVPMHQGSLAA 504  
|||||  
QY 414 GPCGCCSSCLNIGSKGKSSYCEPTLLKANKKQNFSEVMVPVKASGSSHCRAGQTGVC 473  
|||||  
Db 505 GPCGCCSSCLNIGSKGKSSYCEPTLLKANKKQNFSEVMVPVKASGSSHCRAGQTGVC 564  
|||||  
QY 474 LIM 476  
|||||  
Db 565 LIM 567  
  
RESULT 4  
AAU19443  
ID AAU19443 standard; Protein; 298 AA.  
XX  
XX AAU19443;  
XX  
DT 04-DEC-2001 (first entry)  
XX  
DE Human diagnostic and therapeutic polypeptide (DITHP) #29.  
XX  
KW Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;  
KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;  
KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;  
KW respiratory disorder.  
XX  
OS Homo sapiens.  
XX





```
XX SQ Sequence 355 AA;
Query Match 49.6%; Score 1246; DB 22; Length 355;
Best Local Similarity 73.1%; Pred. No. 9.4e-105;
Matches 237; Conservative 40; Mismatches 45; Indels 2; Gaps 2;

QY 1 MGRKEEDCCSWKQTNRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPA 59
Db 1 MARENGESSWKKQAEIDIKKIFEFKETLGTGAFSEVFLAEKATGKLFVAKCIPKALK 60

QY 60 FRDSSLENEIAVLKKIKHENIVTLEDIYESTHYLYVMQLVSGGELFDRILRGVYTEKD 119
Db 61 GKESSENEIAVLKKIKHENIVALEDIYESPNNHLYVMQLVSGGELFDRIVERGFTYTKD 120

QY 120 ASLVIQQVLSAVKYLHENGIVHRDLKPNLLYLTPPEENSKIMITDFGLSKMEQNG-IMST 178
Db 121 ASTLIROVLDAVYVYLRMGIVHRDLKPNLLYYSQDESKIMISDFGLSKMEKGDMVST 180

QY 179 ACGTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPPPYEETSKLPEKIKEGYEF 238
Db 181 ACGTPGYVAPEVLAQKPYSKAVDCWSIGVIAIYILLCGYPPPYDENDSKLPEQILKAEYEF 240

QY 239 ESPFWDDISESAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIOKN 298
Db 241 DSPYWDIDSDSAKDFIRNLMEKDPNKRYTCEQAARHPWIAAGDTALNKNIHESVSAQIRKN 300

QY 299 FAKSKWROAFNAAVVHMRKLHM 322
Db 301 FAKSKWROAFNATAVVRHMRKLHL 324

RESULT 7
ABB08178
ID ABB08178 standard; Protein: 355 AA.
XX
AC ABB08178;
XX
DT 23-SEP-2002 (first entry)
XX
DE Human Cam kinase I 39.05 polypeptide.
XX
KW Human; Cam kinase I 39.05; nervous disease; arrhythmia; tumour;
KW enzyme.
XX
OS Homo sapiens.
XX
PN CN1333357-A.
XX
PD 30-JAN-2002.
XX
PF 07-JUL-2000; 2000CN-0119405.
XX
PR 07-JUL-2000; 2000CN-0119405.
XX
PA (SHAN-) SHANGHAI BIODOOR GENE DEV CO LTD.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2002-305611/35.
DR N-PSDB; ABL60905.
XX
PT Novel polypeptide human Cam kinase I 39.05 and polynucleotide for
PT encoding said polypeptide.
XX
PS Claim 1; Page 27-28 (disclosure); 35pp; Chinese.
XX
CC The present invention discloses a novel polypeptide-human Cam kinase I
CC 39.05, polynucleotide for coding the polypeptide and method for producing
CC this polypeptide by DNA recombination technology. The polypeptide is
CC useful for treating diseases such as nervous disease, arrhythmia, tumour
CC and growth development disturbance disease. The present sequence
CC represents the human Cam kinase I 39.05 polypeptide.
```

```
XX SQ Sequence 355 AA;
Query Match 49.6%; Score 1246; DB 23; Length 355;
Best Local Similarity 73.1%; Pred. No. 9.4e-105;
Matches 237; Conservative 40; Mismatches 45; Indels 2; Gaps 2;

QY 1 MGRKEEDCCSWKQTNRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPA 59
Db 1 MARENGESSWKKQAEIDIKKIFEFKETLGTGAFSEVFLAEKATGKLFVAKCIPKALK 60

QY 60 FRDSSLENEIAVLKKIKHENIVTLEDIYESTHYLYVMQLVSGGELFDRILRGVYTEKD 119
Db 61 GKESSENEIAVLKKIKHENIVALEDIYESPNNHLYVMQLVSGGELFDRIVERGFTYTKD 120

QY 120 ASLVIQQVLSAVKYLHENGIVHRDLKPNLLYLTPPEENSKIMITDFGLSKMEQNG-IMST 178
Db 121 ASTLIROVLDAVYVYLRMGIVHRDLKPNLLYYSQDESKIMISDFGLSKMEKGDMVST 180

QY 179 ACGTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPPPYEETSKLPEKIKEGYEF 238
Db 181 ACGTPGYVAPEVLAQKPYSKAVDCWSIGVIAIYILLCGYPPPYDENDSKLPEQILKAEYEF 240

QY 239 ESPFWDDISESAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIOKN 298
Db 241 DSPYWDIDSDSAKDFIRNLMEKDPNKRYTCEQAARHPWIAAGDTALNKNIHESVSAQIRKN 300

QY 299 FAKSKWROAFNAAVVHMRKLHM 322
Db 301 FAKSKWROAFNATAVVRHMRKLHL 324

RESULT 8
AAE11768
ID AAE11768 standard; Protein: 357 AA.
XX
AC AAE11768;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human kinase (PKIN)-2 protein.
XX
KW Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
KW cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
KW myasthenia gravis; cirrhosis; cataract; growth and development disorder;
KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
KW asthma; obesity; restorative therapy; cytostatic; immunomodulatory;
KW antimicrobial; cardiovascular; antiinflammatory; vaccine.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..40
FT Domain /label= Signal_peptide
FT Domain 24..270
FT Domain /note= "Protein kinase domain"
FT Domain 23..279
FT Protein /note= "Eukaryotic protein kinase domain"
FT Protein 41..357
FT Protein /label= Human_mature_PKIN-2_protein
XX
PW WO200181555-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12992.
XX
PR 20-APR-2000; 2000US-199021P.
PR 28-APR-2000; 2000US-200226P.
PR 05-MAY-2000; 2000US-202339P.
```



QY 1 MGRKEEDDCSSWKKQTTNIRKTFIFMEVLGSAFSEVFLVKQRLTGKLFALKCI-KKSPA 59  
Db 1 MARENGESSWKKQAEIDIKKIFEFKETLGTGAFFSEVLAEEKATGKLFVAVKCIKKALK 60  
QY 60 FRDSSLENEIAVLKIKHENIVTLEDIYESTHYLYVMQVSGGELFDRILRGVYTEKD 119  
Db 61 GKSSLENEIAVLKIKHENIVALEDIYESPNNHLYVMQVSGGELFDRIVEKGFYTEKD 120  
QY 120 ASLVIOQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKIMITDFGLSKMEQNG-IMST 178  
Db 121 ASTLIQRLVDVAVYLHRMGIVHRDLKPNLLYYSQDEESKIMISDFGLSKMEGKGDVWMT 180  
QY 179 ACCTPGYVAPEVLAQPKYSKAVDCWSIGVITYILLGYPFPYETESKLFKIKEGYEF 238  
Db 181 ACCTPGYVAPEVLAQPKYSKAVDCWSIGVIAYILLGYPFPYDENDSKLFEQILKAEYEF 240  
QY 239 ESPFWDDISAKDFICHLEKDPNERVTCESKALSHPWIDGNTALHRDLYPSVSIQOKN 298  
Db 241 DSPYWDIDSDSAKDFIRNMEKDPNKRYTCEQAARHPWIAGDTALNKNIHESVSAQIRKN 300  
QY 299 FAKSKWRQAFNAAVVHMRKLHM 322  
Db 301 FAKSKWRQAFNATAVVRHMRKLHL 324  
RESULT 10  
AAB84359  
ID AAB84359 standard; Protein; 385 AA.  
AC AAB84359;  
DT 22-AUG-2001 (first entry)  
DE Amino acid sequence of a human kinase polypeptide.  
KW Human; kinase; human disease; human disorder.  
OS Homo sapiens.  
XX WO20012435-A2.  
PN 14-JUN-2001.  
PD 07-DEC-2000; 2000WO-US3240.  
PF 07-DEC-1999; 99US-0169428.  
PR (LEXI-) LEXICON GENETICS INC.  
PA Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;  
PI AbuIn A, Sands AT;  
XX WPI; 2001-381667/40.  
DR N-PSDB; AAB25118.  
XX Novel isolated human kinase polynucleotide that shares structural  
PT similarity with animal kinases including calcium/calmodulin-dependent  
PT protein kinases and serine/threonine protein kinases, useful in  
PT therapeutics -  
XX Claim 3; Page 29-30; 32pp; English.  
XX The present sequence represents a kinase polypeptide. The kinase  
CC polynucleotides and polypeptides are useful in therapeutic, diagnostic  
CC and pharmacogenetic applications. They are useful for the detection of  
CC mutant kinases, or inappropriately expressed kinases for the diagnosis  
CC of a disease or disorder. They are useful for screening for drugs (or  
CC high throughput screening of combinatorial libraries) effective in the  
CC treatment of symptomatic or phenotypic manifestations of that disease  
CC or disorder. The polynucleotide sequence is useful as a source of  
CC probes and primers, which can be used to screen libraries, isolate  
CC clones, and prepare cloning and sequencing templates.  
XX

SQ Sequence 385 AA;  
Query Match 49.6%; Score 1246; DB 22; Length 385;  
Best Local Similarity 73.1%; Pred. No. 1.1e-104;  
Matches 237; Conservative 40; Mismatches 45; Indels 2; Gaps 2;  
QY 1 MGRKEEDDCSSWKKQTTNIRKTFIFMEVLGSAFSEVFLVKQRLTGKLFALKCI-KKSPA 59  
Db 1 MARENGESSWKKQAEIDIKKIFEFKETLGTGAFFSEVLAEEKATGKLFVAVKCIKKALK 60  
QY 60 FRDSSLENEIAVLKIKHENIVTLEDIYESTHYLYVMQVSGGELFDRILRGVYTEKD 119  
Db 61 GKSSLENEIAVLKIKHENIVALEDIYESPNNHLYVMQVSGGELFDRIVEKGFYTEKD 120  
QY 120 ASLVIOQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKIMITDFGLSKMEQNG-IMST 178  
Db 121 ASTLIQRLVDVAVYLHRMGIVHRDLKPNLLYYSQDEESKIMISDFGLSKMEGKGDVWMT 180  
QY 179 ACCTPGYVAPEVLAQPKYSKAVDCWSIGVITYILLGYPFPYETESKLFKIKEGYEF 238  
Db 181 ACCTPGYVAPEVLAQPKYSKAVDCWSIGVIAYILLGYPFPYDENDSKLFEQILKAEYEF 240  
QY 239 ESPFWDDISAKDFICHLEKDPNERVTCESKALSHPWIDGNTALHRDLYPSVSIQOKN 298  
Db 241 DSPYWDIDSDSAKDFIRNMEKDPNKRYTCEQAARHPWIAGDTALNKNIHESVSAQIRKN 300  
QY 299 FAKSKWRQAFNAAVVHMRKLHM 322  
Db 301 FAKSKWRQAFNATAVVRHMRKLHL 324  
RESULT 11  
AAU03508  
ID AAU03508 standard; Protein; 357 AA.  
AC AAU03508;  
XX 12-SEP-2001 (first entry)  
DT Human protein kinase #8.  
DE Human; protein kinase; PTK; STK; cancer; cardiovascular disease;  
KW metabolic disorder; immune related disease; neurological disorder;  
KW neurodegenerative disorder; inflammatory disorder; infectious disease;  
KW reproductive disorder.  
XX Homo sapiens.  
XX WO200138503-A2.  
XX 31-MAY-2001.  
XX 22-NOV-2000; 2000WO-US32085.  
XX 24-NOV-1999; 99US-0167482.  
XX (SUGE-) SUGEN INC.  
XX Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;  
PI Flanagan P, Clary D;  
XX WPI; 2001-343950/36.  
DR N-PSDB; AAS06708.  
XX Nucleic acids encoding human kinase polypeptides, useful for preventing  
CC diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
CC neuronal-associated diseases, and microbial infections -  
PS Claim 7; Figure 2; 433pp; English.  
XX AAU03501-AAU03557 represent novel human protein kinases #1-57. The  
CC novel protein kinases have been identified as members of the tyrosine  
CC or serine/threonine kinase (PTK and STK) families. The polynucleotides



```
XX PR 07-DEC-1999; 99US-0169428.
XX PA (LEXI-) LEXICON GENETICS INC.
XX PI Doncho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
XX PI Abulin A, Sands AT;
XX WPI; 2001-381667/40.
XX N-PSDB; AAH25119.
XX Novel isolated human kinase polynucleotide that shares structural
XX similarity with animal kinases including calcium/calmodulin-dependent
XX protein kinases and serine/threonine protein kinases, useful in
XX therapeutics -
XX Claim 4; Page 31; 32pp; English.
XX The present sequence represents a kinase polypeptide. The kinase
XX polynucleotides and polypeptides are useful in therapeutic, diagnostic
XX and pharmacogenic applications. They are useful for the detection of
XX mutant kinases, or inappropriately expressed kinases for the diagnosis
XX of a disease or disorder. They are useful for screening for drugs (or
XX high throughput screening of combinatorial libraries) effective in the
XX treatment of symptomatic or phenotypic manifestations of that disease
XX or disorder. The polynucleotide sequence is useful as a source of
XX probes and primers, which can be used to screen libraries, isolate
XX clones, and prepare cloning and sequencing templates.
XX Sequence 356 AA;
XX Query Match 49.0%; Score 1231.5; DB 22; Length 356;
XX Best Local Similarity 72.8%; Pred. No. 2e-103;
XX Matches 236; Conservative 40; Mismatches 45; Indels 3; Gaps 3;
QY 1 MGRKEEDCSSKKKOTTRIKTFIFMEVLGSGAFSEFLVKQRLTGKLFALKCI-KKSPA 59
DB 1 MARENGESSSSKKQAEIDIKIFEKELGTGAFSEVLAEEKATGKLFVAKCIPKRALK 60
QY 60 FRDSSLENEIAVLKIKHENIVTLEDIYESTHYLLVMQLVSGGELFDRILRGVYTEKD 119
DB 61 GRESSIENEIAVLKIKHENIVALEDIYESPNNHLYVMQLVSGGELFDRIVEKGFYTEKD 120
QY 120 ASLVTOQVLAVKYLHENGIVHRDLKPNLLYLPPENSKIMTDFGLSKMEQNG-IMST 178
DB 121 ASLTLRQVLDAVYLLHRRGIVVHRDLKPNLLYYSQDEESKIMISDFGLSKMEGRGDVMT 180
QY 179 ACGTGYVAPEVLAQPKYSKAVDCWSIGVITYILLGPPPEETESKLEKIKEGYEF 238
DB 181 ACGTGYVAPEVLAQPKYSKAVDCWSIGVIAIYLLGPPPYDENDSKLFEQILKAYEF 240
QY 239 ESPFWDDISAKDPFICHLLEKDPNERYTCESKALSHPWIDGNTALHRDIYPSVSLQIQKN 298
DB 241 DSPYWDIDISAKDPFIRNLMEKDPNRYTCQQAARHPWI-ADTALNKNIHESVSAQIRKN 299
QY 299 FAKSKWRQAFNAVVHMRKLHM 322
DB 300 FAKSKWRQAFNATAVVRHMRKLHL 323
RESULT 14
RAY68793
ID AAY68793 standard; Protein; 389 AA.
XX AC AAY68793;
XX DT 16-MAY-2000 (first entry)
XX DE Amino acid sequence of a human phosphorylation effector PHSP-25.
XX KW Human; phosphorylation effector; PHSP; proliferative disorder;
XX immune disorder; neuronal disorder.
XX
```

```
OS Homo sapiens.
XX Key Location/Qualifiers
XX Modified-site 31 /note= "potential phosphorylation site"
XX Modified-site 56 /note= "potential phosphorylation site"
XX Domain 73..311 /note= "protein kinase catalytic domain"
XX Modified-site 96 /note= "potential phosphorylation site"
XX Modified-site 134 /note= "potential phosphorylation site"
XX Modified-site 148 /note= "potential phosphorylation site"
XX Modified-site 149 /note= "potential phosphorylation site"
XX Domain 152..208 /note= "potential phosphorylation site"
XX Modified-site 165 /note= "PTK core domain"
XX Region 172..184 /note= "potential phosphorylation site"
XX Modified-site 186 /note= "STK core catalytic motif"
XX Modified-site 201 /note= "potential phosphorylation site"
XX Modified-site 257 /note= "potential phosphorylation site"
XX Modified-site 283 /note= "potential glycosylation site"
XX Modified-site 301 /note= "potential phosphorylation site"
XX Modified-site 343 /note= "potential phosphorylation site"
XX Modified-site 358 /note= "potential glycosylation site"
XX Modified-site 364 /note= "potential phosphorylation site"
XX Modified-site 375 /note= "potential glycosylation site"
XX W0200006728-A2.
XX 10-FEB-2000.
XX 28-JUL-1999; 99WO-US17132.
XX 28-JUL-1998; 98US-0123494.
XX 14-SEP-1998; 98US-0152814.
XX 03-OCT-1998; 98US-0173482.
XX 03-NOV-1998; 98US-0106889.
XX 19-NOV-1998; 98US-0109093.
XX 22-DEC-1998; 98US-0113796.
XX 12-JAN-1999; 99US-0173482.
XX 12-JAN-1999; 99US-0229005.
XX (INCY-) INCYTE PHARM INC.
XX Hillman JL, Lal P, Tang YT, Corley NC, Guegler KJ, Baughn MR;
XX Patterson C, Bandman O, Au-Young J, Gorgone GA, Yue H, Azimzal Y;
XX Reddy R, Lu DAM, Shih LL;
XX WPI; 2000-183125/16.
XX N-PSDB; AA246162.
XX New human phosphorylation effectors useful for the diagnosis, treatment
XX and prevention of proliferative, immune and neuronal disorders -
XX Claim 1; Page 112-113; 142pp; English.
XX AAY68769-95 and AAY68797-99 represent human phosphorylation effectors
XX (PHSP), designated PHSP1-PHSP31 (the protein sequence for PHSP28 is not
```





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 17:23:21 ; Search time 3882 Seconds  
(without alignments)  
10208.747 Million cell updates/sec

Title: US-09-960-643-1

Perfect score: 2447

Sequence: 1 tggagtgaggactcaagcag.....ttttctctataaaaaaaaa 2447

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	787	32.2	1126	BM547443	BM547443 AGENCOURT
2	762	31.1	809	BI821474	BI821474 603038366
3	710	29.0	740	BI824483	BI824483 603038855
4	684	28.0	1048	BI084101	BI084101 602869466
5	665	27.2	1129	BM807335	BM807335 AGENCOURT
c 6	656	26.8	966	BI084897	BI084897 602869466

7	650	26.6	731	13	BI753035	BI753035 603025844
8	636	26.0	742	13	BI818261	BI818261 603032510
9	577	23.6	812	13	BI772626	BI772626 603060879
10	551	22.5	1068	14	BM921532	BM921532 AGENCOURT
c 11	479	19.6	479	14	BQ102407	BQ102407 121c07.x
12	473	19.3	824	12	BG715920	BG715920 602676667
13	450	18.4	501	9	AL134342	AL134342 DKF2P5470
c 14	426	17.4	484	10	AW016039	AW016039 UT-H-B10p
c 15	410	16.8	410	9	AA838372	AA838372 of29a11.s
16	408	16.7	573	14	BQ086330	BQ086330 121c07.y
17	284	11.6	554	13	BI834635	BI834635 603090418
18	275	11.2	656	13	BI667965	BI667965 603292877
19	241	9.8	315	9	AA351937	AA351937 EST59855
20	139	5.7	537	17	AQ265651	AQ265651 CITBT-E1-
21	109	4.5	510	17	AQ041427	AQ041427 CIT-HSP-2
22	89	3.6	89	12	BG170935	BG170935 602323912
23	80	3.3	386	14	R05661	R05661 yf87d09.r1
24	71	2.9	818	12	BG293660	BG293660 602390529
25	71	2.9	905	14	BQ934044	BQ934044 AGENCOURT
c 26	70	2.9	273	12	BF950790	BF950790 RC3-NN019
27	68	2.8	188	17	AQ078706	AQ078706 CIT-HSP-2
c 28	59	2.4	306	9	AI713924	AI713924 UI-R-AC1-
c 29	59	2.4	439	10	AW522224	AW522224 UI-R-B00-
c 30	59	2.4	451	10	AW254051	AW254051 UI-R-BJ0-
c 31	59	2.4	555	10	AW251224	AW251224 UI-R-BJ0-
c 32	57	2.3	445	9	AI215131	AI215131 qp41f02.x
33	56	2.3	974	12	BG173948	BG173948 602334031
34	50	2.0	288	12	BF524803	BF524803 UI-R-AC1-
35	36	1.5	375	10	BB871015	BB871015 BB871015
c 36	35	1.4	118	10	AW122501	AW122501 UI-M-BH2
c 37	35	1.4	204	10	AW521699	AW521699 UI-R-B00-
c 38	35	1.4	375	12	BE944047	BE944047 UI-M-BH3-
c 39	35	1.4	377	10	AW492153	AW492153 UI-M-BH3-
c 40	35	1.4	425	12	BF599348	BF599348 263103.MA
c 41	35	1.4	433	9	AI850410	AI850410 UI-M-BG1-
c 42	35	1.4	503	12	BG381458	BG381458 UI-R-CTO-
c 43	32	1.3	262	10	AW955600	AW955600 EST367670
c 44	32	1.3	286	14	T93936	T93936 ye06e06.r1
c 45	32	1.3	314	9	AA365669	AA365669 EST76508

#### ALIGNMENTS

RESULT 1  
BM547443 1126 bp mRNA linear EST 20-FEB-2002  
LOCUS AGENCOURT\_6507688 NIH\_MGC\_125 Homo sapiens cDNA clone IMAGE:5724450  
DEFINITION 5', mRNA sequence.  
ACCESSION BM547443  
VERSION BM547443.1 GI:18781213  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1126)  
AUTHORS NIH-MGC http://mgc.ncbi.nlm.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Invitrogen  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL12713 row: 1 column: 19  
High quality sequence stop: 623.  
Location/Qualifiers  
1. .1126  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5724450"  
/clone\_lib="NIH\_MGC\_125"  
/lab\_host="DH10B"  
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;  
Site\_1: EcorV (destroyed); Site\_2: NotI; RNA source pool  
of three ovaries, from females ranging in age from 38 to  
49 yo. Library is oligo-dT primed and directionally cloned  
(EcorV site is destroyed upon cloning). Average insert  
size 2.1 kb, insert size range 1-3.5 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 036."

BASE COUNT 299 a 291 c 286 g 244 t 6 others  
ORIGIN

Query Match 32.2%; Score 787; DB 13; Length 1126;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 837; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGAGTGGAGCTCAAGCAGGATTTCTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACTC 61  
Db 1 GGAGTGGAGCTCAAGCAGGATTTCTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACTC 60  
Qy 62 TGGAGGCAATGGGTGCGAAAGAAAGAGATGACTGCAGTTCTTGGGAAGAAACAGACCACCA 121  
Db 61 TGGAGGCAATGGGTGCGAAAGAAAGAGATGACTGCAGTTCTTGGGAAGAAACAGACCACCA 120  
Qy 122 ACATCCGGAACCTTCATTTTATGGAGTGTCTGGGATCAGAGCTTCTCAGAAGTTT 181  
Db 121 ACATCCGGAACCTTCATTTTATGGAGTGTCTGGGATCAGAGCTTCTCAGAAGTTT 180  
Qy 182 TCCTGGTGAACRAAGACTGACTGGGAAGCTTTCTGCTGAGTGCATCAAGAGTCAC 241  
Db 181 TCCTGGTGAACRAAGACTGACTGGGAAGCTTTCTGCTGAGTGCATCAAGAGTCAC 240  
Qy 242 CTGCTCTCCGGGACAGCAGCTGGAGAAATGAGATGCTGTGTGTTGAAAAAGATCAAGCATG 301  
Db 241 CTGCTCTCCGGGACAGCAGCTGGAGAAATGAGATGCTGTGTGTTGAAAAAGATCAAGCATG 300  
Qy 302 AAACATTTGACCTCGGAGACATCTATGAGAGCACCACCTACTACTCTGGTTCATCG 361  
Db 301 AAACATTTGACCTCGGAGACATCTATGAGAGCACCACCTACTACTCTGGTTCATCG 360  
Qy 362 AGCTTTGTTCTGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAGA 421  
Db 361 AGCTTTGTTCTGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAGA 420  
Qy 422 AGGATGCCAGTCTGGTGATCCAGAGGTCTTGTGGGAGTGAATACCTACATGAGAATG 481  
Db 421 AGGATGCCAGTCTGGTGATCCAGAGGTCTTGTGGGAGTGAATACCTACATGAGAATG 480  
Qy 482 GCATCTGCCACAGAGACTTAAAGCCGAAACCTGCTTACCTTACCTGGAAGAGAACT 541  
Db 481 GCATCTGCCACAGAGACTTAAAGCCGAAACCTGCTTACCTTACCTGGAAGAGAACT 540  
Qy 542 CTAAGATCATGATCACTGACTTTGGTCTGTCCAGATGGAACAGATGGCATCATGTCGA 601  
Db 541 CTAAGATCATGATCACTGACTTTGGTCTGTCCAGATGGAACAGATGGCATCATGTCGA 600  
Qy 602 CTGCTGTGGGACCCAGGCTACGTGGCTCCAGAAGTGTCTGGCCGACAGAACCCCTACAGCA 661  
Db 601 CTGCTGTGGGACCCAGGCTACGTGGCTCCAGAAGTGTCTGGCCGACAGAACCCCTACAGCA 660  
Qy 662 AGGCTGTGGATGCTGTGTCATCGGCGTCATCACCTACATATTTGCTCTGTGGATACCCCC 721  
Db 661 AGGCTGTGGATGCTGTGTCATCGGCGTCATCACCTACATATTTGCTCTGTGGATACCCCC 720  
Qy 722 CGTTCTATGAGAAACCGAGTCTAAGCTTTTCGAGAGATCAAGGAGGCTACTATGAGT 781  
Db 721 CATTTCTATGAGAAACCGAGTCTAAGCTTTTCGAGAGATCAAGGAGGCTACTATGAGT 780

Qy 782 TTGAGTCTCCATCTCTGGGATGACATTTCTAGTCAGCCAGGACTTTATTTCGCCACTT 839  
Db 781 TTGAGTCTCCATCTCTGGGATGACATTTCTAGTCAGCCAGGACTTTATTTCGCCACTT 838

RESULT 2  
BI821474

LOCUS BI821474 809 bp mRNA linear EST 04-OCT-2001  
DEFINITION 603038366F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5179336 5',  
mRNA sequence.

ACCESSION BI821474

VERSION BI821474.1 GI:15933024

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 809)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L14M11447 row: c column: 17

High quality sequence stop: 808.

FEATURES

Location/Qualifiers

1..809

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5179336"

/clone\_lib="NIH\_MGC\_115"

/lab\_host="DH10B"

/note="Organ: pooled brain, lung, testis; Vector:

pCMV-SPORT6; Site\_1: NotI; Site\_2: EcorV (destroyed); RNA

source anonymous pool of 6 male brains, age range 23-27; 1

male lung, age 27; and 1 male testis, age 69. Library is

oligo-dT primed and directionally cloned [EcorV site is

destroyed upon cloning]. Average insert size 1.8 kb,

insert size range 1-3 kb. Library is normalized and

enriched for full-length clones and was constructed by C.

Gruber (Invitrogen). Research Genetics tracking code

021. Note: this is a NIH-MGC Library."

BASE COUNT 212 a 192 c 218 g 187 t

ORIGIN

Query Match 31.1%; Score 762; DB 13; Length 809;

Best Local Similarity 100.0%; Pred. No. 5.1e-306;

Matches 762; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGGAGTGGAGCTCAAGCAGGATTTCTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT 60

Db 6 TGGAGTGGAGCTCAAGCAGGATTTCTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT 65

Qy 61 CTGGAGGCAATGGGTGCGAAAGAAAGAGATGACTGCAGTTCTCTGGGAAGAAACAGACCAC 120

Db 66 CTGGAGGCAATGGGTGCGAAAGAAAGAGATGACTGCAGTTCTCTGGGAAGAAACAGACCAC 125

Qy 121 AACATCCGGAACCTTCATTTTATGGAAGTGTCTGGATCAGAGCTTCTCAGAAGTT 180

Db 126 AACATCCGGAACCTTCATTTTATGGAAGTGTCTGGATCAGAGCTTCTCAGAAGTT 185

Qy 181 TTCCTGTGGAACAAAGACTGACTGGGAAGCTTTGCTCTGGAAGTGCATCAAGAGTCA 240

Db 186 TTCCTGTGGAACAAAGACTGACTGGGAAGCTTTGCTCTGGAAGTGCATCAAGAGTCA 245

Qy 241 CCTGCTTCCGGGACAGCAGCCTCGAGNATGAGATTGCTGTGTTGAAAAGATCAAGCAT 300

[illegible]

```

RESULT 3
BI824483
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

740 bp mRNA linear EST 04-OCT-2001
603038855F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179957 5',
mRNA sequence.
BI824483
BI824483.1 GI:15936033
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 740)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11448 row: m column: 14
High quality sequence stop: 709.
Location/Qualifiers
1..740
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5179957"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/notes="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27;

```

male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH-MGC Library." 196 a 196 c 167 t

BASE COUNT	196 a	181 c	196 g	167 t	ORIGIN
Query Match	29.04%	Score 710;	DB 13;	Length 740;	
Best Local Similarity	100.0%;	Pred. No. 2.2e-284;			
Matches 710;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	TGGAGTGGGAGCTCAAGCAGAGATTCTTCCCGAGTCCCTGGCATCCTCAGAGCTTCAACT	60		
Db	1	TGGAGTGGGAGCTCAAGCAGAGATTCTTCCCGAGTCCCTGGCATCCTCAGAGCTTCAACT	60		
Qy	61	CTGAGGCAATGGGTGCAAGAGAAAGATGACTGCAGTTCTCTGGAAGAAACAGCACACC	120		
Db	61	CTGAGGCAATGGGTGCAAGAGAAAGATGACTGCAGTTCTCTGGAAGAAACAGCACACC	120		
Qy	121	AACATCCGGAAAACTTCATTTTATGGAAGTGTGGGATCAGAGCTTTCTCAGAAGTT	180		
Db	121	AACATCCGGAAAACTTCATTTTATGGAAGTGTGGGATCAGAGCTTTCTCAGAAGTT	180		
Qy	181	TTCTGTGTGAAGCAAAAGACTGACTGGGAAGCTCTTTGCTCTGAAAGTGCATCAAGAAGTCA	240		
Db	181	TTCTGTGTGAAGCAAAAGACTGACTGGGAAGCTCTTTGCTCTGAAAGTGCATCAAGAAGTCA	240		
Qy	241	CCTGCTTCCGGGACAGCAGCTCGAGAAATGAGATTCGTGTGTTGAAAGATCAAGCAT	300		
Db	241	CCTGCTTCCGGGACAGCAGCTCGAGAAATGAGATTCGTGTGTTGAAAGATCAAGCAT	300		
Qy	301	GAAACATTGTGACCCCTGGAGGACATCTATGAGAGCACACCCACTACTCTGGTCATG	360		
Db	301	GAAACATTGTGACCCCTGGAGGACATCTATGAGAGCACACCCACTACTCTGGTCATG	360		
Qy	361	CAGCTTGTCTGTGGGAGCTCTTTGACCGGATCCCTGGAGCGGGTGTCTACACAGAG	420		
Db	361	CAGCTTGTCTGTGGGAGCTCTTTGACCGGATCCCTGGAGCGGGTGTCTACACAGAG	420		
Qy	421	AAGATGCCAGTCTGGTGCATCCAGCAGCTCTTGTGCGCAGTGAATACCTACATGAGAA	480		
Db	421	AAGATGCCAGTCTGGTGCATCCAGCAGCTCTTGTGCGCAGTGAATACCTACATGAGAA	480		
Qy	481	GGCATCGTCCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC	540		
Db	481	GGCATCGTCCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC	540		
Qy	541	TCTAAGATCATGATCAGTTCAGTTCCTGCTGTCAGATGGAACAAATGGCATCATGTCC	600		
Db	541	TCTAAGATCATGATCAGTTCAGTTCCTGCTGTCAGATGGAACAAATGGCATCATGTCC	600		
Qy	601	ACTGCTGTGGACCCAGGCTACGTGGCTCCAGAAAGTCTGGCCCGCAAAACCTTACAGC	660		
Db	601	ACTGCTGTGGACCCAGGCTACGTGGCTCCAGAAAGTCTGGCCCGCAAAACCTTACAGC	660		
Qy	661	AAGCTGTGGATTCGTGTGTCATCCAGCTTACCTACATATTCCTCTG	710		
Db	661	AAGCTGTGGATTCGTGTGTCATCCAGCTTACCTACATATTCCTCTG	710		
RESULT 4					
LOCUS	BI084101				
DEFINITION	002869466F1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5013856 5',	1048 bp	mRNA	linear	EST 20-JUN-2001
ACCESSION	BI084101				
VERSION	BI084101.1	GI:14502431			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1048)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: ATCC

cDNA Library Preparation: CLONETECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: L1CM1820 row: d column: 17  
High quality sequence stop: 853.

#### FEATURES

source location/Qualifiers  
1..1048  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5013856"  
/clone\_lib="NIH\_MGC\_102"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: salivary gland; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed  
by Ling Hong in the laboratory of Gerald M. Rubin  
(University of California, Berkeley) using ZP-cDNA  
synthesis kit (Stratagene) and Superscript II RT (Life  
Technologies). Note: this is a NIH\_MGC Library."  
BASE COUNT 231 a 298 c 264 g 254 t 1 others  
ORIGIN

Query Match 28.0%; Score 684; DB 13; Length 1048;

Best Local Similarity 100.0%; Pred. No. 1.2e-273;  
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1425 GGTGATGTTAAAGCCAGTGGCAGCTCCCACTCCGCGGAGGCGAGCTGGAGT 1484

DB 2 GGTGATGTTAAAGCCAGTGGCAGCTCCCACTCCGCGGAGGCGAGCTGGAGT 61

QY 1485 CTGCTCATATGATGTTCTGAGCGCTGTGCTATGTCAGTGCATTTTCAGGAGACAT 1544

DB 62 CTGCTCATATGATGTTCTGAGCGCTGTGCTATGTCAGTGCATTTTCAGGAGACAT 121

QY 1545 ATTCAACTCTCTGCTCTTCCAAACCTGGTCTATCCGCGAGAGGAGGAGGAGGAGC 1604

DB 122 ATTCAACTCTCTGCTCTTCCAAACCTGGTCTATCCGCGAGAGGAGGAGGAGGAGC 181

QY 1605 AAGTGGAGCAGGCTTAGCAGGAGCAGTTCCTGGCCAGAACCAAGCCTGCTGCCAGG 1664

DB 182 AAGTGGAGCAGGCTTAGCAGGAGCAGTTCCTGGCCAGAACCAAGCCTGCTGCCAGG 241

QY 1665 GGCAGCCCTCATAGAGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1724

DB 242 GGCAGCCCTCATAGAGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 301

QY 1725 TGTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1784

DB 302 TGTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 361

QY 1785 CCCACACCTAGTCCGCTGGCTGTGTCAGTGTACGTAGTACGTACGTACGTACGTACGT 1844

DB 362 CCCACACCTAGTCCGCTGGCTGTGTCAGTGTACGTAGTACGTACGTACGTACGTACGT 421

QY 1845 GTGCTGTTGCTGTAAGAGCTTAATGGGCTGGCCAGGCTGTGTCACCTTCTCCAGCAA 1904

DB 422 GTGCTGTTGCTGTAAGAGCTTAATGGGCTGGCCAGGCTGTGTCACCTTCTCCAGCAA 481

QY 1905 AGCCATATGGAGCATACCCAGACTCCCACTCTGTCACACACTCACTCCCACTCTCAAG 1964

Db 482 AGCCATATGGAGCATACCCAGACTCCCACTCTGTCACACACTCACTCCCACTCTCAAG 541  
QY 1965 CTCTCAACCTCTTGCCAGAGATTGGCTCATTAATGTGCTGCCCTGCCATCTGCATGAAT 2024  
Db 542 CTCTCAACCTCTTGCCAGAGATTGGCTCATTAATGTGCTGCCCTGCCATCTGCATGAAT 601  
QY 2025 GACAGGCAGCTCCCATGGTGTGCTGCTGAGCTCTTCAAGTCTTAATCCTTAACCTCC 2084  
Db 602 GACAGGCAGCTCCCATGGTGTGCTGCTGAGCTCTTCAAGTCTTAATCCTTAACCTCC 661  
QY 2085 AGGATTAGTCCCAAGTCGCTGA 2108  
Db 662 AGGATTAGTCCCAAGTCGCTGA 685

#### RESULT 5

BM807335 1129 bp mRNA linear EST 05-MAR-2002  
LOCUS AGENCOURT\_6575096 NIH\_MGC\_124 Homo sapiens cDNA clone IMAGE:5732298  
5', mRNA sequence.

ACCESSION BM807335

VERSION BM807335.1 GI:19124158

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1129)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Invitrogen

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM12734 row: c column: 19

High quality sequence start: 11

High quality sequence stop: 651.

FEATURES Location/Qualifiers

1..1129

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5732298"

/clone\_lib="NIH\_MGC\_124"

/tissue\_type="hippocampus"

/lab\_host="DH10B"

/note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: EcoRV

(destroyed); Site\_2: NotI; RNA source male hippocampus,

age 27. Library is oligo-dT primed and directionally

cloned (EcoRV site is destroyed upon cloning). Average

insert size 1.4 kb, insert size range 0.9-4 kb. Library is

normalized and enriched for full-length clones and was

constructed by C. Gruber (Invitrogen). Research Genetics

tracking code 012."

BASE COUNT 287 a 308 c 294 g 240 t

ORIGIN

Query Match 27.2%; Score 665; DB 14; Length 1129;

Best Local Similarity 100.0%; Pred. No. 9.5e-266;

Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGAGTGGAGCTCAAGCAGGATTCTTCCCGAGTCCCTGGCATCTCCAGAGCTTCAACT 60

Db 36 TGGAGTGGAGCTCAAGCAGGATTCTTCCCGAGTCCCTGGCATCTCCAGAGCTTCAACT 95

QY 61 CTGAGGCAATGGGTCGAAAGAGATGACTGCAAGTTCCTGGAAAGAACAGACCACC 120

|||||

Db 96 CTGGAGCAATGGTGCAGAAAGGAAGATGACTCAGCTTCCTGGAAGAAACAGACCACC 155  
Qy 121 AACATCCGGAACACCTTCATTTTATGGAAGTGTGGATCAGGAGCTTCTCAGAAGTT 180  
Db 156 AACATCCGGAACACCTTCATTTTATGGAAGTGTGGATCAGGAGCTTCTCAGAAGTT 215  
Qy 181 TTCTCTGGTGAAGCAAGACTGACTGGGAAGCTTTTGTCTGCTGAAGTGCATCAAGAAGTCA 240  
Db 216 TTCTCTGGTGAAGCAAGACTGACTGGGAAGCTTTTGTCTGCTGAAGTGCATCAAGAAGTCA 275  
Qy 241 CCTGCTTCCGGGACAGCAGCCTGGAGATGAGATGCTGTGTTGAAAAAGATCAAGCAT 300  
Db 276 CCTGCTTCCGGGACAGCAGCCTGGAGATGAGATGCTGTGTTGAAAAAGATCAAGCAT 335  
Qy 301 GAAACATTTGTGACCTGGAGACATCTATGAGACACACCCACTACTCTGTCATG 360  
Db 336 GAAACATTTGTGACCTGGAGACATCTATGAGACACACCCACTACTCTGTCATG 395  
Qy 361 CAGCTTGTCTTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAG 420  
Db 396 CAGCTTGTCTTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAG 455  
Qy 421 AAGGATGCCAGTCTGGTGATCCAGCAGGTCTTTGTCGGCAGTGAATACTACATGAGAAT 480  
Db 456 AAGGATGCCAGTCTGGTGATCCAGCAGGTCTTTGTCGGCAGTGAATACTACATGAGAAT 515  
Qy 481 GGCATGCTCCAGAGACTTAAGCCCGAAGACCTGTCTTACCTTACCCCTGAAGAGAAC 540  
Db 516 GGCATGCTCCAGAGACTTAAGCCCGAAGACCTGTCTTACCTTACCCCTGAAGAGAAC 575  
Qy 541 TCTAAGATCATGATCACTGACTTGTGCTGTCCTCAAGATGGACACAAATGGCATCATGCC 600  
Db 576 TCTAAGATCATGATCACTGACTTGTGCTGTCCTCAAGATGGACACAAATGGCATCATGCC 635  
Qy 601 ACTGCTGTGGGACCCAGGCTACGTGGCTCCAGAAAGTCTGGCCAGCAAAACCCCTACAGC 660  
Db 636 ACTGCTGTGGGACCCAGGCTACGTGGCTCCAGAAAGTCTGGCCAGCAAAACCCCTACAGC 695  
Qy 661 AAGGC 665  
Db 696 AAGGC 700

RESULT 6  
BI084897/c  
LOCUS 602869466r1 NIH\_MGC\_102 Homo sapiens cDNA clone IMAGE:5013856 3',  
DEFINITION mRNA sequence.  
ACCESSION BI084897  
VERSION BI084897.1 GI:14503227  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 966)  
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: CLONETECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM1820 row: d column: 17  
High quality sequence stop: 832.  
Location/Qualifiers  
1..966  
/organism="Homo sapiens"

/db\_xref="taxon:9606"  
/clone="IMAGE:5013856"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/notes="Organ: salivary gland; Vector: pOTB7; Site\_1: xhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCAGAG(G). Library constructed  
by Ling Hong in the laboratory of Gerald M. Rubin  
(University of California, Berkeley) using ZAP-cDNA  
synthesis kit (Stratagene) and Superscript II RT (Life  
Technologies). Note: this is a NIH\_MGC Library."  
BASE COUNT 219 a 238 c 302 g 207 t  
ORIGIN  
Query Match 26.8%; Score 656; DB 13; Length 966;  
Best Local Similarity 99.9%; Pred. No. 5.5e-262;  
Matches 776; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Qy 1589 GGGAGGAGGAGAGAGAGTGGAGCAGGGCTTAGCAGGAGCAGTTCCTGGCCAGAGCAC 1648  
Db 797 GGGAGGAGGAGAGAGTGGAGCAGGGCTTAGCAGGAGCAGTTCCTGGCCAGAGCAC 738  
Qy 1649 CAGCCTGCTGCCAGCGGGCAGCCCTCATAGGAGGCCAGGAGGAGCCCAAGGCGTA 1708  
Db 737 CAGCCTGCTGCCAGCGGGCAGCCCTCATAGGAGGCCAGGAGGAGCCCAAGGCGTA 678  
Qy 1709 GAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACAGCTTCCAGGCTCCCTCG 1768  
Db 677 GAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACAGCTTCCAGGCTCCCTCG 618  
Qy 1769 ACCTGCTGCTCATATGCCACACCCCTACGTGCCGTGGCTCTGTGCACTGTACAGTATA 1828  
Db 617 ACCTGCTGCTCATATGCCACACCCCTACGTGCCGTGGCTCTGTGCACTGTACAGTATA 558  
Qy 1829 GCTCTGCTGGGCTGTGCTGTTGCTGTAAGAGCTTAATGGGCTGGCCAGGCTCTGT 1888  
Db 557 GCTCTGCTGGGCTGTGCTGTTGCTGTAAGAGCTTAATGGGCTGGCCAGGCTCTGT 498  
Qy 1889 CACCTTCTCCAGCAAGCCATATGAGCAGCTTACCCAGAGCTCCCACTCTGCACACTC 1948  
Db 497 CACCTTCTCCAGCAAGCCATATGAGCAGCTTACCCAGAGCTCCCACTCTGCACACTC 438  
Qy 1949 ACTCCAGCTCTAAGCCTCCACCTCTTGGCCAGATTGGGCTCATTAATGTCGTCCT 2008  
Db 437 ACTCCAGCTCTAAGCCTCCACCTCTTGGCCAGATTGGGCTCATTAATGTCGTCCT 378  
Qy 2009 GCCCATCTGCATGAATGACAGGAGCTCCCATGCTGGCTGTGAGCTCTTCAAGT 2068  
Db 377 GCCCATCTGCATGAATGACAGGAGCTCCCATGCTGGCTGTGAGCTCTTCAAGT 318  
Qy 2069 TCTAATCCTTAACCTCCAGGATTAGTCCCAAGTGCCTGAGACCCAGCCAGCAGCTTCT 2128  
Db 317 TCTAATCCTTAACCTCCAGGATTAGTCCCAAGTGCCTGAGACCCAGCCAGCAGCTTCT 258  
Qy 2129 GGCCCTTCTCCTGCCTCAATCTAAAAGCAGTGCACACCCCTCCAAAGTGAATAGAAAG 2188  
Db 257 GGCCCTTCTCCTGCCTCAATCTAAAAGCAGTGCACACCCCTCCAAAGTGAATAGAAAG 199  
Qy 2189 AAGTTCATGAGTAAAGGCTGCAAGGAATTCCTATCTTGGCCACATCTCCCTGGCACAC 2248  
Db 198 AAGTTCATGAGTAAAGGCTGCAAGGAATTCCTATCTTGGCCACATCTCCCTGGCACAC 139  
Qy 2249 ACCCAATGGAGTTAACTTGAAGTGTGACTATTTAATGCTGCGCAGGAGTTCTTAATCCT 2308  
Db 138 ACCCAATGGAGTTAACTTGAAGTGTGACTATTTAATGCTGCGCAGGAGTTCTTAATCCT 79  
Qy 2309 GCCTCTGTTCCTTTTCTCTCTTTGAAAGTCCAGCAGCAGTTCCTGTCCTTCCCA 2365  
Db 78 GCCTCTGTTCCTTTTCTCTCTTTGAAAGTCCAGCAGCAGTTCCTGTCCTTCCCA 22

```
RESULT 7
BI753035
LOCUS
DEFINITION
603025844F1 NIH_MGC_114 Homo sapiens cdna clone IMAGE:5196323 5',
mRNA sequence.
ACCESSION
BI753035
VERSION
BI753035.1 GI:15744613
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 731)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1491 row: g column: 12
High quality sequence stop: 729.
FEATURES
source
1. .731
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5196323"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/Note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH MGC Library."
BASE COUNT
190 a 182 c 191 g 168 t
ORIGIN
Query Match 26.6%; Score 650; DB 13; Length 731;
Best Local Similarity 100.0%; Pred. No. 1.9e-259;
Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 73 GGTCAAGGAAGAGATGACTGCTCGGAGTCTTCTCAGAGTTCCTCAGAGTTCCTCGTGAAG 132
Db 73 GGTCAAGGAAGAGATGACTGCTCGGAGTTCCTCAGAGTTCCTCAGAGTTCCTCGTGAAG 132
QY 133 ACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAGTTCCTCAGAGTTCCTCGTGAAG 192
Db 133 ACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAGTTCCTCAGAGTTCCTCGTGAAG 192
QY 193 CAAGAGCTGACTGGGAAGCTTTTGCTGTGAAGTGCATCAAGAAAGTCAACCTGCTCCCGG 252
Db 193 CAAGAGCTGACTGGGAAGCTTTTGCTGTGAAGTGCATCAAGAAAGTCAACCTGCTCCCGG 252
QY 253 GACAGCAGCTGGAGATGAGATGCTGTGTGAAAGATCAAGCATGAACATTTGTG 312
Db 253 GACAGCAGCTGGAGATGAGATGCTGTGTGAAAGATCAAGCATGAACATTTGTG 312
QY 313 ACCCTGGAGGACATCTATGAGAGCACCACCACCTACTACCTGGTCACTGAGCTGTTTCT 372
Db 313 ACCCTGGAGGACATCTATGAGAGCACCACCACCTACTACCTGGTCACTGAGCTGTTTCT 372
QY 373 GGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGGATGCCAGT 432
Db 373 GGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGGATGCCAGT 432
```

```
QY 433 CTGGTGATCCAGCAGGTCTTGTGCGCAGTGAATAACCTACATGAGATGSCATCGTCCAC 492
|||||
Db 433 CTGGTGATCCAGCAGGTCTTGTGCGCAGTGAATAACCTACATGAGATGSCATCGTCCAC 492
|||||
QY 493 AGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAACTCTTAAGATCATG 552
|||||
Db 493 AGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAACTCTTAAGATCATG 552
|||||
QY 553 ATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGGCATATGTCCTGCTGCTGGG 612
|||||
Db 553 ATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGGCATATGTCCTGCTGCTGGG 612
|||||
QY 613 ACCCCAGGCTACGTGGCTCCAGAAAGTCTGGCCCGAGAAACCCCTACAGCAAGGCTGTGGAT 672
|||||
Db 613 ACCCCAGGCTACGTGGCTCCAGAAAGTCTGGCCCGAGAAACCCCTACAGCAAGGCTGTGGAT 672
|||||
QY 673 TGCTGGTCCATCGGCTCATCACTACATATTCCTCTGTGGATACCCCCC 722
|||||
Db 673 TGCTGGTCCATCGGCTCATCACTACATATTCCTCTGTGGATACCCCCC 722
|||||
RESULT 8
BI818261
LOCUS
DEFINITION
603032510F1 NIH_MGC_115 Homo sapiens cdna clone IMAGE:5173587 5',
mRNA sequence.
ACCESSION
BI818261
VERSION
BI818261.1 GI:15928724
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 742)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1432 row: d column: 04
High quality sequence stop: 742.
FEATURES
source
1. .742
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5173587"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site: 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH MGC Library."
BASE COUNT
197 a 180 c 197 g 168 t
ORIGIN
Query Match 26.0%; Score 636; DB 13; Length 742;
Best Local Similarity 99.9%; Pred. No. 1.2e-253;
Matches 686; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 17 GCAGGATCTTCCCGAGTCCCTGGCATCTCAGAGCTTCAACTCTGGAGCAATGGGTC 76
Db 24 GCAGGATCTTCCCGAGTCCCTGGCATCTCAGAGCTTCAACTCTGGAGCAATGGGTC 83
QY 77 GAAAGGAAGAAGATGACTGCAGTTCCTGGAAGAAACAGACCACCAACATCCGGAAGCT 136
Db 84 GAAAGGAAGAAGATGACTGCAGTTCCTGGAAGAAACAGACCACCAACATCCGGAAGCT 143
QY 137 TCATTTTATGGAAGTCTGGATCAGGAGCTTCTCAGAAGTTTCTGGTGAAGCAAA 196
Db 144 TCATTTTATGGAAGTCTGGATCAGGAGCTTCTCAGAAGTTTCTGGTGAAGCAAA 203
QY 197 GACTGACTGGAAGCTCTTTGCTGCAAGTGCATCAAGAAAGTCACTGCCCTCCGGGACA 256
Db 204 GACTGACTGGAAGCTCTTTGCTGCAAGTGCATCAAGAAAGTCACTGCCCTCCGGGACA 263
QY 257 GCAGCTGGAGATGAGATTGCTGTGTTGAAAAAGATCAAGCATGAAACATTTGTGACCC 316
Db 264 GCAGCTGGAGATGAGATTGCTGTGTTGAAAAAGATCAAGCATGAAACATTTGTGACCC 323
QY 317 TGGAGACATCTATGAGAGCACCACCCTACTACTGCTGTCAGCTTGTCTGGTG 376
Db 324 TGGAGACATCTATGAGAGCACCACCCTACTACTGCTGTCAGCTTGTCTGGTG 383
QY 377 GGGAGCTCTTTGACCGGATCTTGGAGCGGGTGTCTACACAGAGAAGATGCCAGTCTGG 436
Db 384 GGGAGCTCTTTGACCGGATCTTGGAGCGGGTGTCTACACAGAGAAGATGCCAGTCTGG 443
QY 437 TGATCCAGCAGGTCTTTGTCGGCAGTGAATACCTACATGAGAAATGGCATCGTCCACAG 496
Db 444 TGATCCAGCAGGTCTTTGTCGGCAGTGAATACCTACATGAGAAATGGCATCGTCCACAG 503
QY 497 ACTTAAGCCCGAAACCTTGTACCTTACCCCTGAGAGAACTCTTAAGATCATGATCA 556
Db 504 ACTTAAGCCCGAAACCTTGTACCTTACCCCTGAGAGAACTCTTAAGATCATGATCA 563
QY 557 CTGACTTTGGTCTGTCCAAAGTGGACAGATGGCATATGCCACTGCCCTGTGGGACCC 616
Db 564 CTGACTTTGGTCTGTCCAAAGTGGACAGATGGCATATGCCACTGCCCTGTGGGACCC 623
QY 617 CAGGCTACGTGGCTCCAGAAAGTGTGGCCCGAGAAACCTTACAGCAAGGCTGTGGATTGCT 676
Db 624 CAGGCTACGTGGCTCCAGAAAGTGTGGCCCGAGAAACCTTACAGCAAGGCTGTGGATTGCT 683
QY 677 GFTCCATCGGCTCATCCTACATAT 703
Db 684 GFTCCATCGGCTCATCCTACATAT 710

RESULT 9
BI772626
LOCUS 812 bp mRNA linear EST 25-SEP-2001
DEFINITION 603060879F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5210299 5',
mRNA sequence.
ACCESSION BI772626
VERSION BI772626.1 GI:15764204
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 812)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
```

```
http://image.llnl.gov
Plate: LUAM1527 row: m column: 20
High quality sequence start: 2
High quality sequence stop: 808.
FEATURES
Location/Qualifiers
source
1..812
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5210299"
/lab_lib="NIH_MGC_122"
/lab_post="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed, and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT 217 a 186 c 221 g 188 t
ORIGIN
Query Match 23.6%; Score 577; DB 13; Length 812;
Best Local Similarity 99.8%; Pred. No. 3.9e-229;
Matches 627; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TGGAGTGGGAGCTCAAGCAGGATTTCTCCGAGTCCCTGGCATCTCAGAGCTTCAACT 60
Db 9 TGGAGTGGGAGCTCAAGCAGGATTTCTCCGAGTCCCTGGCATCTCAGAGCTTCAACT 68
QY 61 CTGAGGCAATGGGTGCAAGAGAGATGACTGCGAGTTCTCTGGAGAAACAGACACC 120
Db 69 CTGAGGCAATGGGTGCAAGAGAGATGACTGCGAGTTCTCTGGAGAAACAGACACC 128
QY 121 AACATCCGGAAACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTCTCAGAAGTT 180
Db 129 AACATCCGGAAACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTCTCAGAAGTT 188
QY 181 TTCCTGCTGGAAGCAAGACGACTCTGGGAAGCTCTTCTCTGAAGTGCATCAAGAACTCA 240
Db 189 TTCCTGCTGGAAGCAAGACGACTGCTGGGAAGCTCTTCTCTGAAGTGCATCAAGAACTCA 248
QY 241 CCTGCCTTCGGGACAGCAGCCTGGAGAAATGAGATTGCTGTGTGAAAAGATCAAGCAT 300
Db 249 CCTGCCTTCGGGACAGCAGCCTGGAGAAATGAGATTGCTGTGTGAAAAGATCAAGCAT 308
QY 301 GAAACATTTGACCCCTGGAGGACATCTATGAGAGCACCAACCTACTACTCTGGTCATG 360
Db 309 GAAACATTTGACCCCTGGAGGACATCTATGAGAGCACCAACCTACTACTCTGGTCATG 368
QY 361 CAGCTTCTTCTGCTGGTGGGAGCTCTTTTACCGGATCCTGGAGCGGGGTGTCTACACAG 420
Db 369 CAGCTTCTTCTGCTGGTGGGAGCTCTTTTACCGGATCCTGGAGCGGGGTGTCTACACAG 428
QY 421 AAGGATCCAGTCTGCTGATCCAGCAGGCTTGTTCGCGCATGAAATACCTACATGAGAT 480
Db 429 AAGGATCCAGTCTGCTGATCCAGCAGGCTTGTTCGCGCATGAAATACCTACATGAGAT 488
QY 481 GGCATCTGCACAGAGACTTTAAAGCCCGAAAACCTGCTTTACCTTACCTCGAAGAGAAC 540
Db 489 GGCATCTGCACAGAGACTTTAAAGCCCGAAAACCTGCTTTACCTTACCTCGAAGAGAAC 548
QY 541 TCTAAGATCATGATCATGACTGCTTGTCTCCAAAGATGGAACAGAAATGGCATCATGTCC 600
Db 549 TCTAAGATCATGATCATGACTGCTTGTCTCCAAAGATGGAACAGAAATGGCATCATGTCC 608
QY 601 ACTGCTCTGGGACCCAGGCTACGTGG 628
Db 609 ACTGCTCTGGGACCCAGGCTACGTGG 636
```

```
RESULT 10
BM921532
LOCUS
DEFINITION
AGENCOURT_6708041 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5753010
5', mRNA sequence.
ACCESSION
BM921532
VERSION
BM921532.1 GI:19371911
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1068)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12788 row: b column: 19
High quality sequence stop: 592.
FEATURES
Location/Qualifiers
1..1068
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5753010"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT 279 a 281 c 260 g 248 t
ORIGIN
Query Match 22.5%; Score 551; DB 14; Length 1068;
Best Local Similarity 99.8%; Pred. No. 2.3e-218;
Matches 601; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TGGAGTGGGAGCTCAAGCAGGATTCTCCCGAGTCCTCGCATCCTCGAAGCTTCAACT 60
|||||
Db 14 TGGAGTGGGAGCTCAAGCAGGATTCTCCCGAGTCCTCGCATCCTCGAAGCTTCAACT 73
|||||
QY 61 CTGGAGCAATGGTCCGAAGGAAGATGACTGCAGTCTCTGGAAGAAACAGACCACC 120
|||||
Db 74 CTGGAGCAATGGTCCGAAGGAAGATGACTGCAGTCTCTGGAAGAAACAGACCACC 133
|||||
QY 121 AACATCCGGAACCTTCATTTTATGGAAGTCCTGGGATCAGGAGCTTCTCAGAAGTT 180
|||||
Db 134 AACATCCGGAACCTTCATTTTATGGAAGTCCTGGGATCAGGAGCTTCTCAGAAGTT 193
|||||
QY 181 TTCTCTGTGAAGCAAAAGACTGACTGGGAAGCTTTTGCTCTGAAGTCATCAAGAAGTCA 240
|||||
Db 194 TTCTCTGTGAAGCAAAAGACTGACTGGGAAGCTTTTGCTCTGAAGTCATCAAGAAGTCA 253
|||||
QY 241 CCTGCCTCCGGACAGCAGCCCTGGAGATGAGATTGCTGTGTGTAAGAAAGATCAAGCAT 300
|||||
Db 254 CCTGCCTCCGGACAGCAGCCCTGGAGATGAGATTGCTGTGTGTAAGAAAGATCAAGCAT 313
|||||
QY 301 GAAAAATGTGACCTGGAGGACATCTATGAGAGCACCACCACCTACTACCTGGTCATG 360
|||||
```

```
Db 314 GAAAAATGTGACCTGGAGGACATCTATGAGAGCACCACCACCTACTACCTGGTCATG 373
QY 361 CAGCTTGTCTTGTGGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAG 420
|||||
Db 374 CAGCTTGTCTTGTGGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAG 433
|||||
QY 421 AAGGATGCCAGTCTGGTGATCCAGCAGGTCTTGTCCGAGTGAATACCTACATGAGAAT 480
|||||
Db 434 AACGATGCCAGTCTGGTGATCCAGCAGGTCTTGTCCGAGTGAATACCTACATGAGAAT 493
|||||
QY 481 GGCATCTGCACACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCTCCCTGAAGAGAAC 540
|||||
Db 494 GGCATCTGCACACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCTCCCTGAAGAGAAC 553
|||||
QY 541 TCTAAGATCATGATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGCGATCATGTC 600
|||||
Db 554 TCTAAGATCATGATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGCGATCATGTC 613
|||||
QY 601 AC 602
||
Db 614 AC 615

RESULT 11
BQ102407/c
LOCUS
DEFINITION
BQ102407.1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:6135181 3', mRNA sequence.
ACCESSION
BQ102407
VERSION
BQ102407.1 GI:20135391
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 479)
REFERENCE
Melton D., Brown J., Kenty G., Permutt, A., Lee C., Kaestner, K.,
Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blustain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas
M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.,
Jackson, Y., and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other_ESTs: ij2ic07.y1
COMMENT
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@bioh.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center This clone is
available royalty-free through LLNL; please contact the IMAGE
consortium (info@image.llnl.gov) for further information
Seq primer: -40UP from Gibco
High quality sequence stop: 448.
FEATURES
Location/Qualifiers
1..479
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6135181"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
Site_2: Sal 1; Starting library constructed using
SuperScript plasmid library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
```

fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an Ecot of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT	136 a	89 c	147 g	107 t
ORIGIN				
Query Match	19.6%; Score 479; DB 14; Length 479;			
Best Local Similarity	100.0%; Pred. No. 2.5e-188;			
Matches 479; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
QY 1960	TCAAGCCTCAACCTCTTGGCCAGATTGGGCTCATTAATGTGCTGCTGCCCATCTGCA 2019			
Db	TCAAGCCTCAACCTCTTGGCCAGATTGGGCTCATTAATGTGCTGCTGCCCATCTGCA 420			
QY 2020	TGAATGACAGGACGCTCCCATGCTGCTGCTGAGACTCTTCAAGTTCTAAATCCTTA 2079			
Db	TGAATGACAGGACGCTCCCATGCTGCTGCTGAGCTCTTCAAGTTCTAAATCCTTA 360			
QY 2080	ACTCCAGGATTAGCTCCCAAGTSCGCTGAGACCCAGCCAGCACACTTCTGCCCTTCTCC 2139			
Db	ACTCCAGGATTAGCTCCCAAGTSCGCTGAGACCCAGCCAGCACACTTCTGCCCTTCTCC 300			
QY 2140	CTGCCTCAATCTAAAAGCAGTGCACACCCCTCCAAAGTGGAAATAGAAAGATTCTATGAG 2199			
Db	CTGCCTCAATCTAAAAGCAGTGCACACCCCTCCAAAGTGGAAATAGAAAGATTCTATGAG 240			
QY 2200	TAAGGGCTGCAAGGAATCTTATCTGCGCCACATGCTCTCCGTGCACACCCCAATGGAG 2259			
Db	TAAGGGCTGCAAGGAATCTTATCTGCGCCACATGCTCTCCGTGCACACCCCAATGGAG 180			
QY 2260	TTAACTTGAAGTCTGACATTTTAAATGCTGCGAGGAGTTCTAATCCTGCTCTGTTCC 2319			
Db	TTAACTTGAAGTCTGACATTTTAAATGCTGCGAGGAGTTCTAATCCTGCTCTGTTCC 120			
QY 2320	CTTTCTCTCTTGAAGTCCAGCACACCATCTTGTCTCTCCCGAGTTTCTCGCCCTC 2379			
Db	CTTTCTCTCTTGAAGTCCAGCACACCATCTTGTCTCTCCCGAGTTTCTCGCCCTC 60			
QY 2380	CACCCCTCCAGCTTCATGCTCAGTGTGCTTAATAAATGGACATATTTTCTCTAA 2438			
Db	CACCCCTCCAGCTTCATGCTCAGTGTGCTTAATAAATGGACATATTTTCTCTAA 1			

RESULT 12	
LOCUS BG715920	
DEFINITION BG715920 824 bp mRNA linear EST 08-MAY-2001	
602676667F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4799224 5', mRNA sequence.	
ACCESSION BG715920	
VERSION BG715920.1	
KEYWORDS GI:13995107	
SOURCE EST.	
ORGANISM human.	
REFERENCE Homo sapiens	
1 (bases 1 to 824)	
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE NIH-MGC http://mgc.nci.nih.gov/.	
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)	
COMMENT Unpublished (1999)	
Contact: Robert Strausberg, Ph.D.	
Email: cgapbs-re@mail.nih.gov	
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.	
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)	
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
DNA Sequencing by: Incyte Genomics, Inc.	

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LIAM0688 row: e column: 17  
High quality sequence stop: 766.

FEATURES	
source	1. .824
Location/Qualifiers	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="IMAGE:4799224"	
/clone_lib="NIH_MGC_96"	
/tissue_type="hypothalamus"	
/lab_host="DH10B"	
/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."	
BASE COUNT	223 a 194 c 219 g 188 t
ORIGIN	

Query Match	19.3%; Score 473; DB 12; Length 824;
Best Local Similarity	99.5%; Pred. No. 6.6e-186;
Matches 623; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
QY 1	TGGAGTGGGAGCTCAAGCAGGATTTCTCCGAGTCCCTGGCATCTCAGAGCTTCAACT 60
Db	TGGAGTGGGAGCTCAAGCAGGATTTCTCCGAGTCCCTGGCATCTCAGAGCTTCAACT 97
QY 61	CTGAGGCAATGGGTGGAAGGAAGAGATGACTGCAAGTTCCTTGGAAAGAAACAGACCACC 120
Db	CTGAGGCAATGGGTGGAAGGAAGAGATGACTGCAAGTTCCTTGGAAAGAAACAGACCACC 157
QY 121	AACATCCGGAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAGTT 180
Db	AACATCCGGAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAGTT 217
QY 181	TTCTGTGTGAAGCAAGAACTGACTGGGAAGCTCTTTGTCTGAAAGTGCATCAAGAAATCA 240
Db	TTCTGTGTGAAGCAAGAACTGACTGGGAAGCTCTTTGTCTGAAAGTGCATCAAGAAATCA 277
QY 241	CCTGCTTTCGGGACAGCAGCCTTGAGAAATGAGATTCGTGTGTTGAAAGATCAAGCAT 300
Db	CCTGCTTTCGGGACAGCAGCCTTGAGAAATGAGATTCGTGTGTTGAAAGATCAAGCAT 337
QY 301	GAAGACATTTGACCTGGAGGACATCTATGAGAGCACCACCTACTACTCTGTCATG 360
Db	GAAGACATTTGACCTGGAGGACATCTATGAGAGCACCACCTACTACTCTGTCATG 397
QY 361	CAGCTTGTCTGTGGGAGCTCTTTGACCGGATCTCTGGAGCGGGTGTCTACACAGAG 420
Db	CAGCTTGTCTGTGGGAGCTCTTTGACCGGATCTGGAGCGGGTGTCTACACAGAG 457
QY 421	AAGGATGCCAGTCTGGTGATCCAGAGGCTTTGCGGAGTGAATACCTACATGAGAAT 480
Db	AAGGATGCCAGTCTGGTGATCCAGAGGCTTTGCGGAGTGAATACCTACATGAGAAT 517
QY 481	GGCATCTCCACAGAGACTTAAAGCCCAAAACCTGCTTTTACCTTACCCCTGAAGAGAC 540
Db	GGCATCTCCACAGAGACTTAAAGCCCAAAACCTGCTTTTACCTTACCCCTGAAGAGAC 577
QY 541	TCTAAGATCATGATCAGTCACTGACTTTTGGTCTGTCCAAAGATGGAACAGATGGCATGTCC 600
Db	TCTAAGATCATGATCAGTCACTGACTTTTGGTCTGTCCAAAGATGGAACAGATGGCATGTCC 637
QY 601	ACTGCTGTGGGACCCCAAGGCTACGT 626
Db	ACTGCTGTGGGACCCCAAGGCTACGT 663

RESULT 13	1885	GTGTCACTTCTTCCAGCAAGCCATATGAGCATCTACCCAGACTCCCACTCTGCACAC	1944
AL134342	361	GTGTCACTTCTTCCAGCAAGCCATATGAGCATCTACCCAGACTCCCACTCTGCACAC	420
LOCUS			
DEFINITION			
ACCESSION			
AL134342	1945	ACTCACTCCCACTCTCAAGCCTTCCAACTCTTGGCCAGATTGGCTCATTAATGTCGTT	2004
AL134342	421	ACTCACTCCCACTCTCAAGCCTTCCAACTCTTGGCCAGATTGGCTCATTAATGTCGTT	480
VERSION			
GI:6602529			
EST			
KEYWORDS			
SOURCE			
human.			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
1 (bases 1 to 501)			
Bloecker,H., Boecher,M., Brandt,P., Mewes,W., Weil,B. and Wiemann			
,S.			
EST (Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B. and Wiemann,S.)			
Unpublished (1999)			
Contact: Bloecker H			
MIPS			
Am Klopferspitz 18a D-82152 Martinsried, Germany			
This is the 5' sequence of the clone insert			
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;			
sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.			
s1 sequence also available.			
This clone (DKFp547O044) is available at the RZPD in Berlin.			
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.			
FEATURES			
Location/Qualifiers			
1..501			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="DKFp547O044"			
/clone_lib="547 (synonym: hfbr1)"			
/tissue_type="brain"			
/dev_stage="fetal"			
/lab_host="X1-2blue"			
/note="vector: pAMP1; Site_1: NotI; Site_2: SalI"			
BASE COUNT			
106 a 154 c 132 g 109 t			
ORIGIN			
Query Match			
Best Local Similarity			
99.8%; Pred. No. 2.8e-176; Length 501;			
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy 1525 CTGCAATTTTCAGGAGACATATTCAACTCCCTCTCTTCCAAACCTGGTGTCTATCCGG			1584
Db 1 CTGCAATTTTCAGGAGACATATTCAACTCCCTCTCTTCCAAACCTGGTGTCTATCCGG			60
Qy 1585 CAGAGGGAGGAGGAGCAGAGCAAGTGGAGAGGGCTTAGCAGGAGCAGTTCCTGGCCAGAA			1644
Db 61 CAGAGGGAGGAGGAGCAGAGCAAGTGGAGAGGGCTTAGCAGGAGCAGTTCCTGGCCAGAA			120
Qy 1645 GCACAGCCTGTCGACGCGGGGAGCCCTCATAGGAGGCCAGGAGGAGGCCCAAGG			1704
Db 121 GCACAGCCTGTCGACGCGGGGAGCCCTCATAGGAGGCCAGGAGGAGGCCCAAGG			180
Qy 1705 CGTAAAGCCCTGTTGAAGCTGTGAGCAGGAGGAGCGGTGCCACACAGCTTCCAGGTCTC			1764
Db 181 CGTAAAGCCCTGTTGAAGCTGTGAGCAGGAGGAGCGGTGCCACACAGCTTCCAGGTCTC			240
Qy 1765 CTTGACCTGCTGCTCTATGCCCCACACCCCTACGTCGCGGTCTGTGAGGTACGTA			1824
Db 241 CTTGACCTGCTGCTCTATGCCCCACACCCCTACGTCGCGGTCTGTGAGGTACGTA			300
Qy 1825 GATAGCTCTCCGCTGGGTCCTGCTGTTTGTGCGTAAAGCTTAATGGCTGGCCAGGCT			1884
Db 301 GATAGCTCTCCGCTGGGTCCTGCTGTTTGTGCGTAAAGCTTAATGGCTGGCCAGGCT			360

NCI\_CGAP\_Col0 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE  
Clones 1057416-1061255, 1144584-1145351) The resulting  
subtracted library contained 4 million recombinants.  
Subtraction was performed as previously described [Bonald  
Lennon & Soares (1996): Normalization and Subtraction:  
Two Approaches To Facilitate Gene Discovery. Genome  
Research 6, 791-806.  
TAG\_LIB=NCI\_CGAP\_Brn26  
TAG\_TISSUE=brain  
TAG\_SEQ=ATAGG\*  
BASE COUNT 134 a 142 g 119 t  
ORIGIN

Query Match 17.4%; Score 426; DB 10; Length 484;  
Best Local Similarity 100.0%; Pred. No. 2.6e-166;  
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2013 ATCTGCATGAATGACAGGAGCTCCCATGGTGGTCTGCTGTGAGCTCTTCAAGTTCTA 2072  
|||||  
Db 441 ATCTGCATGAATGACAGGAGCTCCCATGGTGGTCTGCTGTGAGCTCTTCAAGTTCTA 382  
|||||

QY 2073 ATCTTAACTCCAGGATTAGCTCCCAAGTGCCTGAGACCCAGCCAGCACACTTCTGGCC 2132  
|||||  
Db 381 ATCTTAACTCCAGGATTAGCTCCCAAGTGCCTGAGACCCAGCCAGCACACTTCTGGCC 322  
|||||

QY 2133 CTCTCCCTGCTCAATCTAAAGCAGTGCACACCTCCAAAGTGGATAGAAAGT 2192  
|||||  
Db 321 CTCTCCCTGCTCAATCTAAAGCAGTGCACACCTCCAAAGTGGATAGAAAGT 262  
|||||

QY 2193 TCATGATGAGGCTGCAAGGAATCTTATCTCTGGCCACATCTCCTCGTGACACACCC 2252  
|||||  
Db 261 TCATGATGAGGCTGCAAGGAATCTTATCTCTGGCCACATCTCCTCGTGACACACCC 202  
|||||

QY 2253 AATGGAGTTAACTTGGAGTTGACTATTTTAAATGCTGCCAGGAGTTCTAATCTCGCT 2312  
|||||  
Db 201 AATGGAGTTAACTTGGAGTTGACTATTTTAAATGCTGCCAGGAGTTCTAATCTCGCT 142  
|||||

QY 2313 CTGTCTCCCTTTCTCTCTTGAAGTCCAGCACACATCTTGTCTCTCCAGTTTCT 2372  
|||||  
Db 141 CTGTCTCCCTTTCTCTCTTGAAGTCCAGCACACATCTTGTCTCTCCAGTTTCT 82  
|||||

QY 2373 CGCCCTCCACCCCTCCAGCTTCATGCTCAGTGTGTGCTTAATAAAATGGACATATTTT 2432  
|||||  
Db 81 CGCCCTCCACCCCTCCAGCTTCATGCTCAGTGTGTGCTTAATAAAATGGACATATTTT 22  
|||||

QY 2433 CTCTAA 2438  
|||||  
Db 21 CTCTAA 16

RESULT 15  
AA838372/c  
LOCUS  
DEFINITION of29a11.s1 NCI\_CGAP\_Kid6 Homo sapiens cDNA clone IMAGE:1422524,  
mRNA sequence.  
ACCESSION AA838372  
VERSION AA838372.1 GI:2913171  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 410)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaaps-r@mail.nih.gov  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Stratagene, Inc.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Stratagene, Inc.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Insert length: 944 Std Error: 0.00  
Seq primer: -40ml3 fwd. Et from Amersham.

FEATURES  
source  
1..410  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1422524"  
/clone\_lib="NCI\_CGAP\_Kid6"  
/sex="mixed"  
/tissue\_type="kidney tumor"  
/lab\_host="SOLR (kanamycin resistant)"  
/note="Organ: kidney; Vector: Bluescript SK-; Site: 1:  
EcoRI; Site: 2: XhoI; Cloned unidirectionally. Primer:  
Oligo dT. Pooled kidney tumors. 5' adaptor sequence: 5'  
GAATTCGGCAGAG 3' 3' adaptor sequence: 5'  
CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.0 kb."  
BASE COUNT 112 a 79 c 130 g 89 t  
ORIGIN

Query Match 16.8%; Score 410; DB 9; Length 410;  
Best Local Similarity 100.0%; Pred. No. 1.2e-159;  
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1983 GATTGGGCTCAATTAATGTGCTGCCATCTGCCATGAATGACAGGAGTCCCATG 2042  
|||||  
Db 410 GATTGGGCTCAATTAATGTGCTGCCATCTGCCATGAATGACAGGAGTCCCATG 351  
|||||

QY 2043 GTGTCTGCCGTGAGCTCTTCAAGTTCTAATCTTAACTCCAGGATTAGCTCCCAAGT 2102  
|||||  
Db 350 GTGTCTGCCGTGAGCTCTTCAAGTTCTAATCTTAACTCCAGGATTAGCTCCCAAGT 291  
|||||

QY 2103 CGCTGAGACCCAGCAGCACACTTCTGCCCTTCTCCTCAATCTAAAGCAGTGC 2162  
|||||  
Db 290 CGCTGAGACCCAGCAGCACACTTCTGCCCTTCTCCTCAATCTAAAGCAGTGC 231  
|||||

QY 2163 CACACCTCCAAAGTGGAAATAGAAAGTTCAATGAGTAAAGGGCTGCAAGGAATTCCTAT 2222  
|||||  
Db 230 CACACCTCCAAAGTGGAAATAGAAAGTTCAATGAGTAAAGGGCTGCAAGGAATTCCTAT 171  
|||||

QY 2223 CTTGGCCACATGCTCTCGGTGCACACACCCCAATGGAGTTAACTTGAAGTTGACTATTT 2282  
|||||  
Db 170 CTTGGCCACATGCTCTCGGTGCACACACCCCAATGGAGTTAACTTGAAGTTGACTATTT 111  
|||||

QY 2283 TAATGTCTGCCAGGAGTTCTAATCTGCTCTGTTCCCTTTTCTCTCTCTCTCTCTCTCT 2342  
|||||  
Db 110 TAATGTCTGCCAGGAGTTCTAATCTGCTCTGTTCCCTTTTCTCTCTCTCTCTCTCTCT 51  
|||||

QY 2343 CACACCAATCTTGTCTCCCGAGTTTCTCTCGCCCTCCACCCCTCCAGCT 2392  
|||||  
Db 50 CACACCAATCTTGTCTCCCGAGTTTCTCTCGCCCTCCACCCCTCCAGCT 1

Search completed: March 15, 2003, 05:00:43  
Job time : 3894 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 17:25:25 ; Search time 97 Seconds  
(without alignments)  
7736.475 Million cell updates/sec

Title: US-09-960-643-1  
Perfect score: 2447  
Sequence: 1 tggagtggagctcaagcag.....ttttctctataaaaaaaaaa 2447

Scoring table: OLIGO.NUC  
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents.NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCFUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	287	11.7	362	4 US-08-905-223-89	Sequence 89, Appl
2	23	0.9	296	1 US-08-738-367-8	Sequence 8, Appl
3	23	0.9	3380	2 US-09-156-425-1	Sequence 1, Appl
4	22	0.9	755	2 US-08-737-980-1	Sequence 1, Appl
5	21	0.9	1282	2 US-08-878-989-12	Sequence 12, Appl
6	21	0.9	1282	4 US-09-272-796-12	Sequence 12, Appl
7	20	0.8	3116	1 US-08-149-103-2	Sequence 2, Appl
8	20	0.8	3116	1 US-08-451-883-2	Sequence 2, Appl
9	20	0.8	3307	1 US-08-456-647B-5	Sequence 5, Appl
10	20	0.8	3307	2 US-08-237-401A-5	Sequence 5, Appl
11	20	0.8	3785	1 US-08-445-640-9	Sequence 9, Appl
12	20	0.8	3785	3 US-08-170-558-9	Sequence 9, Appl
13	20	0.8	3785	3 US-08-447-314-9	Sequence 9, Appl
14	20	0.8	3785	3 US-08-445-461-9	Sequence 9, Appl
15	19	0.8	1649	2 US-08-466-120-1	Sequence 1, Appl
16	19	0.8	1649	5 PCT-US94-07266-1	Sequence 1, Appl
17	19	0.8	2025	4 US-09-247-155-58	Sequence 58, Appl
18	19	0.8	2573	2 US-08-884-681-2	Sequence 2, Appl
19	19	0.8	2573	4 US-09-258-643-2	Sequence 2, Appl
20	19	0.8	2746	2 US-09-016-000-12	Sequence 12, Appl
21	19	0.8	5827	4 US-09-813-133A-3	Sequence 3, Appl
22	18	0.7	326	1 US-08-700-575-3	Sequence 3, Appl
23	18	0.7	973	3 US-09-013-881-16	Sequence 16, Appl
24	18	0.7	1001	4 US-09-641-638-444	Sequence 444, App
25	18	0.7	1001	4 US-09-641-638-446	Sequence 446, App
26	18	0.7	1283	4 US-09-282-305-11	Sequence 11, Appl
27	18	0.7	1631	4 US-09-051-239A-1	Sequence 1, Appl

c	28	18	0.7	1665	4	US-09-247-155-72	Sequence 72, Appl
	29	18	0.7	1813	4	US-08-123-934A-1	Sequence 1, Appl
	30	18	0.7	1813	5	PCT-US94-10080-1	Sequence 1, Appl
	31	18	0.7	2056	4	US-08-158-735A-1	Sequence 13, Appl
	32	18	0.7	2070	4	US-09-382-256-13	Sequence 13, Appl
	33	18	0.7	2070	4	US-09-395-115-13	Sequence 13, Appl
	34	18	0.7	2070	4	US-08-436-265-13	Sequence 13, Appl
	35	18	0.7	2070	4	US-09-679-187-13	Sequence 11, Appl
	36	18	0.7	2402	4	US-08-462-467B-11	Sequence 11, Appl
	37	18	0.7	2402	4	US-08-156-735A-3	Sequence 3, Appl
c	38	18	0.7	2402	4	US-08-334-179A-11	Sequence 11, Appl
	39	18	0.7	2846	4	US-08-915-795-1	Sequence 11, Appl
	40	18	0.7	5538	2	US-08-231-193A-55	Sequence 55, Appl
	41	18	0.7	5538	3	US-08-486-273A-55	Sequence 55, Appl
	42	18	0.7	5538	4	US-08-940-086A-55	Sequence 55, Appl
	43	18	0.7	5538	4	US-08-940-035A-55	Sequence 55, Appl
	44	18	0.7	5538	4	US-08-935-105A-55	Sequence 55, Appl
	45	18	0.7	5538	4	US-09-648-797-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1  
US-08-905-223-89  
; Sequence 89, Application US/08905223  
; Patent No. 6222029  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste D.  
; APPLICANT: Duclert, Americ  
; APPLICANT: Lacroix, Bruno  
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS  
; NUMBER OF SEQUENCES: 503  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/905,223  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israel, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 89:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 362 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: DOUBLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: CDNA  
; ORIGINAL SOURCE:  
; ORGANISM: Homo Sapiens  
; TISSUE TYPE: Brain  
; FEATURE:  
; NAME/KEY: sig.peptide  
; LOCATION: 87..191  
; IDENTIFICATION METHOD: Von Heijne matrix  
; OTHER INFORMATION: score 5.6  
; OTHER INFORMATION: seq FIFMEVLGSGAFS/EV  
US-08-905-223-89

```
Query Match      11.7%; Score 287; DB 4; Length 362;
Best Local Similarity 99.7%; Pred. NO. 1.7e-127;
Matches 337; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGGAGTGGAGCTCAAGCAGGATTCTTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT 60
Db 18 TGGAGTGGAGCTCAAGCAGGATTCTTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT 77
QY 61 CTGAGGCAATGGGTGGAAGGAAGATGACTGCAGTCTCTGGAAGAAACAGACCACC 120
Db 78 CTGAGGCAATGGGTGGAAGGAAGATGACTGCAGTCTCTGGAAGAAACAGACCACC 137
QY 121 AACATCCGGAACCTTCATTTTATGGAAGTCTGGATCAGGACTTCTCAGAACTT 180
Db 138 AACATCCGGAACCTTCATTTTATGGAAGTCTGGATCAGGACTTCTCAGAACTT 197
QY 181 TTCTGTGTGAAGCAAGACTGACTGGAAGCTCTTTGTCTCTGAAGTGCATCAAGAAGTCA 240
Db 198 TTCTGTGTGAAGCAAGACTGACTGGAAGCTCTTTGTCTCTGAAGTGCATCAAGAAGTCA 257
QY 241 CCTGCTTCGGGACAGCAGCCCTGGAGATGAGATTGCTGTGTTGAAAAAGATCAAGCAT 300
Db 258 CCTGCTTCGGGACAGCAGCCCTGGAGATGAGATTGCTGTGTTGAAAAAGATCAAGCAT 317
QY 301 GAAACACATTGTGACCTCGGAGGACATCTATGAGAGCAC 338
Db 318 GAAACACATTGTGACCTCGGAGGACATCTATGAGAGCAC 355

RESULT 2
US-08-738-367-8
; Sequence 8, Application US/08738367
; Patent No. 5827688
; GENERAL INFORMATION:
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/738,367
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 296 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

; MOLECULE TYPE: cDNA
US-08-738-367-8
Query Match      0.9%; Score 23; DB 1; Length 296;
Best Local Similarity 100.0%; Pred. NO. 0.23;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2425 ATATTTTCTCTAAAAA 2447
Db 268 ATATTTTCTCTAAAAA 290

RESULT 3
US-09-156-425-1
; Sequence 1, Application US/09156425B
; Patent No. 5962671
; GENERAL INFORMATION:
; APPLICANT: Baker, Brenda F.
; APPLICANT: Cowser, Lex M.
; TITLE OF INVENTION: ANTISENSE MODULATION OF FAN EXPRESSION
; FILE REFERENCE: RTS-0009
; CURRENT APPLICATION NUMBER: US/09/156,425B
; CURRENT FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 3380
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(2766)
US-09-156-425-1

Query Match      0.9%; Score 23; DB 2; Length 3380;
Best Local Similarity 100.0%; Pred. NO. 0.23;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2425 ATATTTTCTCTAAAAA 2447
Db 3350 ATATTTTCTCTAAAAA 3372

RESULT 4
US-08-737-980-1
; Sequence 1, Application US/08737980
; Patent No. 5843773
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Apoptosis Regulating Gene
; NUMBER OF SEQUENCES: 2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/737,980
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: KR 1995-6266
; FILING DATE: 24-MAR-1995
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 755 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; DEVELOPMENTAL STAGE: Fetus at 22 weeks gestation
```

; TISSUE TYPE: Liver  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 35..559  
; FEATURE:  
; NAME/KEY: polyA\_signal  
; LOCATION: 713..718  
; US-08-737-980-1

Query Match 0.9%; Score 22; DB 2; Length 755;  
Best Local Similarity 100.0%; Pred. No. 0.66;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2426 TATTTTCTCTAAAAA 2447  
Db 726 TATTTTCTCTAAAAA 747

RESULT 5  
US-08-878-989-12  
; Sequence 12, Application US/08878989  
; Patent No. 5885803  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl G.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/878,989  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0321 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1282 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PROSNOT06  
; CLONE: 827431  
; US-08-878-989-12

Query Match 0.9%; Score 21; DB 2; Length 1282;  
Best Local Similarity 100.0%; Pred. No. 2.1;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 411 CTACACAGAGAGGATGCCAG 431  
Db 521 CTACACAGAGAGGATGCCAG 541

RESULT 6  
US-09-272-796-12  
; Sequence 12, Application US/09272796  
; Patent No. 6207148  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl G.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/272,796  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/878,989  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0321 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1282 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PROSNOT06  
; CLONE: 827431  
; US-09-272-796-12

Query Match 0.9%; Score 21; DB 4; Length 1282;  
Best Local Similarity 100.0%; Pred. No. 2.1;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 411 CTACACAGAGAGGATGCCAG 431  
Db 521 CTACACAGAGAGGATGCCAG 541

RESULT 7  
US-08-149-103-2  
; Sequence 2, Application US/08149103  
; Patent No. 5750367

```
; GENERAL INFORMATION:
; APPLICANT: Lawrence C. B. Chan
; TITLE OF INVENTION: HUMAN AND MOUSE VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN RECEPTORS
; TITLE OF INVENTION: AND METHODS FOR USE OF SUCH
; TITLE OF INVENTION: RECEPTORS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LYON & LYON
; STREET: 611 West Sixth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: IBM MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/149,103
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 204/052
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-149-103-2
;
; Query Match 0.8%; Score 20; DB 1; Length 3116;
; Best Local Similarity 100.0%; Pred. No. 6.2;
; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1309 GGGCCCTGTGGCTGCTC 1328
Db 39 GGGCCCTGTGGCTGCTC 58

RESULT 8
US-08-451-883-2
; Sequence 2, Application US/08451883
; Patent No. 5798209
; GENERAL INFORMATION:
; APPLICANT: Lawrence C. B. Chan
; TITLE OF INVENTION: HUMAN AND MOUSE VERY LOW DENSITY
; TITLE OF INVENTION: LIPOPROTEIN RECEPTORS AND METHODS FOR
; TITLE OF INVENTION: USE OF SUCH RECEPTORS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LYON & LYON
; STREET: 633 West Fifth Street, Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: IBM MS-DOS (Version 6.22)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,883
; FILING DATE: May 26, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/149,103
; FILING DATE: No. 5798209ember 8, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Knight, Matthew W.
; REGISTRATION NUMBER: 36,846
; REFERENCE/DOCKET NUMBER: 212/268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-451-883-2
;
; Query Match 0.8%; Score 20; DB 1; Length 3116;
; Best Local Similarity 100.0%; Pred. No. 6.2;
; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1309 GGGCCCTGTGGCTGCTC 1328
Db 39 GGGCCCTGTGGCTGCTC 58

RESULT 9
US-08-456-647B-5/c
; Sequence 5, Application US/08456647B
; Patent No. 5811516
; GENERAL INFORMATION:
; APPLICANT: Lemke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,647B
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/237,401
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07251/007002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
```

one

INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3307 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
IMMEDIATE SOURCE:  
CLONE: Tyro-3  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 237..2859  
US-08-456-647B-5

Query Match 0.8%; Score 20; DB 1; Length 3307;  
Best Local Similarity 100.0%; Pred. No. 6.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1317 TGGCTGCTGCTCCAGCTGCC 1336  
|||||  
DB 2789 TGGCTGCTGCTCCAGCTGCC 2770

RESULT 10  
US-08-237-401A-5/c  
Sequence 5, Application US/08237401A  
Patent No. 5837448  
GENERAL INFORMATION:  
APPLICANT: Lenke Ph.D. et al., Greg E.  
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/237,401A  
FILING DATE: 02-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/884,486  
FILING DATE: 15-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Halle Ph.D., Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07251/0007001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 678-5070  
TELEFAX: (619) 678-5099  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3307 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
IMMEDIATE SOURCE:  
CLONE: Tyro-3  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 237..2859  
US-08-237-401A-5

Query Match 0.8%; Score 20; DB 2; Length 3307;  
Best Local Similarity 100.0%; Pred. No. 6.2;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1317 TGGCTGCTGCTCCAGCTGCC 1336  
|||||  
DB 2789 TGGCTGCTGCTCCAGCTGCC 2770

RESULT 11  
US-08-445-640-9/c  
Sequence 9, Application US/08445640  
Patent No. 5709858  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J.  
APPLICANT: Mark, Melanie R.  
APPLICANT: Scadden, David T.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Baron, Will F.  
TITLE OF INVENTION: Protein Tyrosine Kinases  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/445,640  
FILING DATE: 22-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/170558  
FILING DATE: 20-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/157563  
FILING DATE: 23-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 854C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3785 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-445-640-9

Query Match 0.8%; Score 20; DB 1; Length 3785;  
Best Local Similarity 100.0%; Pred. No. 6.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1317 TGGCTGCTGCTCCAGCTGCC 1336  
|||||  
DB 2632 TGGCTGCTGCTCCAGCTGCC 2613

RESULT 12  
US-08-170-558-9/c  
Sequence 9, Application US/08170558  
Patent No. 6001621  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J.  
APPLICANT: Mark, Melanie R.

```
; APPLICANT: Scadden, David T.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Protein Tyrosine Kinases
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/170,558
; FILING DATE: 20-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; TELEPHONE: 415/952-9881
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3785 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-170-558-9

Query Match 0.8%; Score 20; DB 3; Length 3785;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1317 TGGCTGCTGCTCCAGCTGCC 1336
Db 2632 TGGCTGCTGCTCCAGCTGCC 2613

RESULT 13
US-08-447-314-9/c
; Sequence 9, Application US/08447314
; Patent No. 6087144
; GENERAL INFORMATION:
; APPLICANT: Scadden, David T.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Protein Tyrosine Kinases
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,314
```

```
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170558
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 854C1D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/952-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3785 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-447-314-9

Query Match 0.8%; Score 20; DB 3; Length 3785;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1317 TGGCTGCTGCTCCAGCTGCC 1336
Db 2632 TGGCTGCTGCTCCAGCTGCC 2613

RESULT 14
US-08-445-461-9/c
; Sequence 9, Application US/08445461
; Patent No. 6096527
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie R.
; APPLICANT: Scadden, David T.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Protein Tyrosine Kinases
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,461
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170558
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 854C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
```





GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:08:41 ; Search time 16 Seconds  
(without alignments)  
1371.239 Million cell updates/sec

Title: US-09-960-643-2

Perfect score: 2513

Sequence: 1 MGRKEEDDCSSWKQTNR.....VKASGSSHRAGQTGVCLIM 476

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2513	100.0	476	9	US-09-935-464-5
2	2362	94.0	460	9	US-09-935-464-3
3	1246	49.6	357	9	US-10-024-036B-2
4	1152.5	45.9	370	9	US-10-142-356-7
5	1152.5	45.9	370	10	US-09-817-181-4
6	1151	45.8	317	9	US-09-935-464-36
7	1016.5	40.4	326	10	US-09-817-181-2
8	944.5	37.6	280	10	US-09-835-788A-16
9	747.5	29.7	501	10	US-09-797-039-2
10	747.5	29.7	501	12	US-10-153-921-2
11	625.5	24.9	765	9	US-10-217-357-4
12	625.5	24.9	765	10	US-09-975-326-4
13	625.5	24.9	766	9	US-09-934-406-2
14	625.5	24.9	766	9	US-10-217-357-2
15	625.5	24.9	766	10	US-09-975-326-2
16	606.5	24.1	385	9	US-10-116-332-2
17	594.5	23.7	545	9	US-10-142-356-6
18	592	23.6	556	12	US-10-096-960-4
19	585.5	23.3	648	9	US-10-024-036B-5

20	585.5	23.3	817	10	US-09-992-481-4	Sequence 4, Appli
21	575.5	22.9	543	10	US-09-740-627-1	Sequence 1, Appli
22	571.5	22.7	565	12	US-10-096-960-2	Sequence 23, Appli
23	563.5	22.4	295	9	US-09-988-462-23	Sequence 16, Appli
24	558	22.2	821	9	US-10-081-119-16	Sequence 11, Appli
25	556.5	22.1	522	10	US-09-740-627-11	Sequence 25, Appli
26	542.5	21.6	463	9	US-09-988-462-25	Sequence 17, Appli
27	538.5	21.4	639	10	US-09-854-731-17	Sequence 5, Appli
28	531	21.1	514	10	US-09-922-138-5	Sequence 2, Appli
29	531	21.1	514	10	US-09-841-683-2	Sequence 5, Appli
30	530	21.1	454	10	US-09-771-161A-238	Sequence 238, App
31	514	20.5	298	10	US-09-858-664A-17	Sequence 17, Appli
32	513	20.4	740	10	US-09-771-161A-265	Sequence 265, App
33	513	20.4	740	10	US-09-771-161A-266	Sequence 266, App
34	508.5	20.2	265	9	US-09-764-868-814	Sequence 814, App
35	508.5	20.2	297	10	US-09-828-313-38	Sequence 38, Appli
36	506	20.1	549	10	US-09-828-313-39	Sequence 39, Appli
37	505.5	20.1	406	10	US-09-771-161A-210	Sequence 210, App
38	502.5	20.0	508	10	US-09-858-664A-18	Sequence 18, Appli
39	498	19.8	370	9	US-10-081-119-34	Sequence 34, Appli
40	496.5	19.8	625	10	US-09-854-731-18	Sequence 18, Appli
41	495.5	19.7	623	10	US-09-854-731-4	Sequence 4, Appli
42	490	19.5	464	9	US-09-988-462-22	Sequence 22, Appli
43	486	19.3	576	10	US-09-854-731-19	Sequence 19, Appli
44	483	19.2	387	10	US-09-771-161A-188	Sequence 188, App
45	480.5	19.1	459	10	US-09-740-627-9	Sequence 9, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-935-464-5  
; Sequence 5, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSY  
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300  
; PRIOR FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 476  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-935-464-5

Query Match	100.0%	Score 2513;	DB 9;	Length 476;
Best Local Similarity	100.0%;	Pred No. 1.3e-181;		
Matches 476;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MGRKEEDDCSSWKQTNRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALCKIKSPAF	60	
Db	1	MGRKEEDDCSSWKQTNRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALCKIKSPAF	60	
Qy	61	RDSLENEIAVLAKKHENIVLTLEDIYESTHYLVQVLVSGGELFDRILRGVYTERDA	120	
Db	61	RDSLENEIAVLAKKHENIVLTLEDIYESTHYLVQVLVSGGELFDRILRGVYTERDA	120	
Qy	121	SLVLIQVLSAVKYLHENGIVHRDLKPKENLIYLTPENSKIMITDFGLSKMEQNGIMSTAC	180	
Db	121	SLVLIQVLSAVKYLHENGIVHRDLKPKENLIYLTPENSKIMITDFGLSKMEQNGIMSTAC	180	
Qy	181	GTPGYVAPEVLAQPKYSKAVDCWSIGVITYILLCGYPPPYEETESKLFPEKIEGYEFES	240	
Db	181	GTPGYVAPEVLAQPKYSKAVDCWSIGVITYILLCGYPPPYEETESKLFPEKIEGYEFES	240	

Qy 241 PFWDISAKDFCHLLEKDPNRYTCEKALSHPWIDGNTALHROIYPSVSLQIQKFA 300  
Db 241 PFWDISAKDFCHLLEKDPNRYTCEKALSHPWIDGNTALHROIYPSVSLQIQKFA 300  
Qy 301 KSKWROAFNAVAHVHMRKLHNNLHSPGVPEVNRPPETQASSTRPSSPEITTEAPV 360  
Db 301 KSKWROAFNAVAHVHMRKLHNNLHSPGVPEVNRPPETQASSTRPSSPEITTEAPV 360  
Qy 361 LDHVSALPALQPCQHGRRPTAPGGRSLNCLVNGSLHSSSLVPMHQSLAAGPGCCGS 420  
Db 361 LDHVSALPALQPCQHGRRPTAPGGRSLNCLVNGSLHSSSLVPMHQSLAAGPGCCGS 420  
Qy 421 SCLNIGSKGKSSYCEPTLLKANKKONFKSEVMPVKASGSSHCRAGOTGVCLIM 476  
Db 421 SCLNIGSKGKSSYCEPTLLKANKKONFKSEVMPVKASGSSHCRAGOTGVCLIM 476  
RESULT 2  
US-09-935-464-3  
; Sequence 3, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA  
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300  
; PRIOR FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 460  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-935-464-3

Query Match 94.0%; Score 2362; DB 9; Length 460;  
Best Local Similarity 100.0%; Pred. No. 3e-170;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MGRKEEDCCSWKKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKSPAF 60  
Db 1 MGRKEEDCCSWKKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKSPAF 60  
Qy 61 RDSSLENEIAVLKKIKHENIVTLEDIYESTHYLVNQLVSGGELFDRILRGVYTEKDA 120  
Db 61 RDSSLENEIAVLKKIKHENIVTLEDIYESTHYLVNQLVSGGELFDRILRGVYTEKDA 120  
Qy 121 SLVTOQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKIMITDGLSKMEQNGIMSTAC 180  
Db 121 SLVTOQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKIMITDGLSKMEQNGIMSTAC 180  
Qy 181 GTPGVYVAPEVLAQPKYSKAVDCWSIGVITVILLCGYPFPFYEETESKLFKEIKGYEYES 240  
Db 181 GTPGVYVAPEVLAQPKYSKAVDCWSIGVITVILLCGYPFPFYEETESKLFKEIKGYEYES 240  
Qy 241 PFWDISAKDFCHLLEKDPNRYTCEKALSHPWIDGNTALHROIYPSVSLQIQKFA 300  
Db 241 PFWDISAKDFCHLLEKDPNRYTCEKALSHPWIDGNTALHROIYPSVSLQIQKFA 300  
Qy 301 KSKWROAFNAVAHVHMRKLHNNLHSPGVPEVNRPPETQASSTRPSSPEITTEAPV 360  
Db 301 KSKWROAFNAVAHVHMRKLHNNLHSPGVPEVNRPPETQASSTRPSSPEITTEAPV 360  
Qy 361 LDHVSALPALQPCQHGRRPTAPGGRSLNCLVNGSLHSSSLVPMHQSLAAGPGCCGS 420  
Db 361 LDHVSALPALQPCQHGRRPTAPGGRSLNCLVNGSLHSSSLVPMHQSLAAGPGCCGS 420

Qy 421 SCLNIGSKGKSSYCEPTLLKANKKQ 447  
Db 421 SCLNIGSKGKSSYCEPTLLKANKKQ 447  
RESULT 3  
US-10-024-036B-2  
; Sequence 2, Application US/10024036B  
; Publication No. US20030028004A1  
; GENERAL INFORMATION:  
; APPLICANT: Bandaru, Rajasekhar  
; TITLE OF INVENTION: 68730 and 69112, Protein Kinase  
; TITLE OF INVENTION: Molecules and Uses Therefor  
; FILE REFERENCE: MPI2000-521P1R(W)  
; CURRENT APPLICATION NUMBER: US/10/024,036B  
; CURRENT FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: 60/258222  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 357  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-024-036B-2  
Query Match 49.6%; Score 1246; DB 9; Length 357;  
Best Local Similarity 73.1%; Pred. No. 2.6e-86;  
Matches 237; Conservative 40; Mismatches 45; Indels 2; Gaps 2;  
Qy 1 MGRKEEDCCSWKKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPA 59  
Db 1 MARENGESSSWKKQOEDIKKIFEFKETGLTGAFSEVLAEEKATGKLFVAVKCIKPKALK 60  
Qy 60 FRDSSLENEIAVLKKIKHENIVTLEDIYESTHYLVNQLVSGGELFDRILRGVYTEKD 119  
Db 61 GKSSLENEIAVLKKIKHENIVALEDIYESPNNHLYLVNQLVSGGELFDRILRGVYTEKD 120  
Qy 120 ASLVIQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKIMITDGLSKMEQNG-IMST 178  
Db 121 ASTLIQVLDVAVYLLHRMGIVHRDLKPNLLYLSQDEESKIMITSDFGLSKMEKGDMVMT 180  
Qy 179 ACGTPGVYVAPEVLAQPKYSKAVDCWSIGVITVILLCGYPFPFYEETESKLFKEIKGYEYEF 238  
Db 181 ACGTPGVYVAPEVLAQPKYSKAVDCWSIGVITVILLCGYPFPFYDENDSKLFEQILKAEYEF 240  
Qy 239 ESPFWDDISAKDFCHLLEKDPNRYTCEKALSHPWIDGNTALHROIYPSVSLQIQKN 298  
Db 241 DSPYWDIDSDSAKDFIRNLMEKDPNKRYTCEQAARHPWIAGDTALNKNIHESVSAQIRKN 300  
Qy 299 FAKSKWROAFNAVAHVHMRKLHM 322  
Db 301 FAKSKWROAFNATAVVRHMRKLHL 324  
RESULT 4  
US-10-142-356-7  
; Sequence 7, Application US/10142356  
; Publication No. US20030036183A1  
; GENERAL INFORMATION:  
; APPLICANT: Boylan, John F.  
; APPLICANT: Bowers, Alex J.  
; TITLE OF INVENTION: Serine-Threonine Kinase Member, h2520-40  
; FILE REFERENCE: 01017/37177A  
; CURRENT APPLICATION NUMBER: US/10/142,356  
; CURRENT FILING DATE: 2002-05-09  
; PRIOR APPLICATION NUMBER: 60/290,276  
; PRIOR FILING DATE: 2001-05-10  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 7  
; LENGTH: 370  
; TYPE: PRT



```

; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-835-788A-16

Query Match          37.6%; Score 944.5; DB 10; Length 280;
Best Local Similarity 77.5%; Pred. No. 9.7e-64;
Matches 176; Conservative 26; Mismatches 24; Indels 1; Gaps 1;

QY  97 MOLVSGGELFDRIILRGVYTEKDSASLYIQVLSAVKYLHENGIVHRDLKPNLLYLTPEE 156
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db   1 MOLVSGGELFDRIIEKGFYTEKDASTLRQVLDVAVYLHRMGIVHRDLKPNLLYSODE 60
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY  157 NSKIMTDFGLSKMEQNG-INSTACGPGYVAPEVLAQPKYSKAVDCWSIGVITYILCG 215
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db   61 ESKIMISDFGLSKMEGKGVNSTATCGTFGYVAPEVLAQPKYSKAVDCWSIGVIAIYILCG 120
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY  216 YPPPYEETESKLFKRIKEGYEFPSPWDDISAKDPICHLEKDPNERYTCEKALSHP 275
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db   121 YPPPYDENSKLFEQILKAEYEFDSYWDIDSADKDFIRNMEKDPNKRYTCEQAARHP 180
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY  276 WIDGNTALHRDIIYPSVSLQIQKNFAKSKWRQAFNAAVVHMRKLHM 322
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db   181 WIAGDTALNKNIHESVSAQIRKNEFAKSKWRQAFNATAVVRHMRKLHL 227
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 9
US-09-797-039-2
; Sequence 2, Application US/09797039
; Patent No. US20020042099A1
; GENERAL INFORMATION:
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 2504, 15977, AND 14760, NOVEL PROTEIN
; TITLE OF INVENTION: KINASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 10448-017001
; CURRENT APPLICATION NUMBER: US/09/797,039
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-797-039-2

Query Match          29.7%; Score 747.5; DB 10; Length 501;
Best Local Similarity 35.9%; Pred. No. 1.5e-48;
Matches 175; Conservative 81; Mismatches 176; Indels 55; Gaps 12;

QY  15 QTTNIRKTFIFMEVLGSGAFSEVFLVKORLTGKLFALKCIKSPAFRD-----SSLENEI 69
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db   16 QPSEVTDYDLGQVIKTEEFCEIFRAKDKTTGKLTCKKKFKQ-----RDGRKVRKAAKNEI 71
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY  70 AVLKKIRHENIVTLEDIESTHYLYVMQLVSGGELFDRILRGVYTEKDSASLYTQOVLS 129
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db   72 GILKMKVHPNIIQLVDVPTKREYFIFLELATGREVDFWILDQGYSERDTSNVVYRQVLE 131
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY  130 AVKYLHENGIVHRDLKPNLLYLTPESKSKMTIDFGLSKMEQNGIMSTACGTPGYVAPE 189
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db   132 AVAYLSLKIIVHRNLKLENLYVNRKLSKIVISDFHLAKLE-NGLIKPECGTPEYLAPE 190
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY  190 VLAQPKYSKAVDCWSIGVITYILCGYPPFYETE-----SKLFKRIKEGYEFESP 241
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db   191 VVGQRQYRPPVDCWAIGVIMYILLSGNPFYEEVEDDYENHDKNLFKILAGDYEFDSP 250
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY  242 FWDDISESAKDFICHLLBKPDNERYTCEKALSHPWIDGNTALHRDIIYPSVSLQIQKNFAK 301
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Db 251 YWDDISQAADLVTRLMVEYEQDQRTAEAEIAISHEWISGNAASDKNIKDGVCQAIEKNFAR 310  
QY 302 SKWRQAFNAAVVHHMRKLIHMNLHSPGVREVENRPPETQASSTRPSPETTITTEAPVL 361  
Db 311 AKWKKAQVRVTTL---MKRL-----RAPE--QSSTAAQASASATDTATPCA 350  
QY 362 ---DHSVALPALTQPCQGRRP-----TAPGRSLNCLVNGSL--HISSSLVPMHQSL 411  
Db 351 AGGATAAAASGATSAPGDAARAASDNVAPADRSATPATDGSVTPATDGS 410  
QY 412 AAGPCGCCSCLN-----IGSKGKSSYCSRPSTLLKANKKQNFSEVWV---PVRASGSSH 464  
Db 411 TPATDGSVPATDRSATPATDGRATPATEESTVPTQSSAMLATKAAATPEPAMAQPDST 470  
QY 465 CRAGQTG 471  
Db 471 APEGATG 477

RESULT 10  
US-10-153-921-2  
; Sequence 2, Application US/10153921  
; Patent No. US20020142430A1  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CL000612DV  
; CURRENT APPLICATION NUMBER: US/10/153,921  
; PRIOR FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/207,281  
; PRIOR FILING DATE: 2000-05-30  
; PRIOR APPLICATION NUMBER: 09/734,030  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 501  
; TYPE: PRT  
; ORGANISM: HOMO SAPIEN  
US-10-153-921-2

Query Match 29.7%; Score 747.5; DB 12; Length 501;  
Best Local Similarity 35.3%; Pred. No. 1.5e-48;  
Matches 173; Conservative 86; Mismatches 170; Indels 61; Gaps 12;

QY 15 QTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKSPAPRD-----SSLENI 69  
Db 16 QPSEVTDRLDQGVKTEEFCEIFRAKDKTTGKLHCKKFKQK---RDGRKVRKAAKNEI 71

QY 70 AVLKIKHENIVTLEDIESTHYILVMQLVSGGELFDRILRGVYTEKDASLVITQOVL 129  
Db 72 GILKWKHPNLIQLVDFVTRKEYFFILELATGREVFDILDQGYYSERDTSNVVRQVLE 131

QY 130 AVKYLHENGIVHRDLKPENLLYLTPENSKIMITDPLGSKMBONGTMSACGTPGVVAPE 189  
Db 132 AVAYLHSLKIVHRDLKLENLVYTRNLRKNSKIVISDFHLAKLE--NLGIKEPCGTPELAPE 190

QY 190 VLAQKPYSAKVDWCWSIGVITYILLCGYPPFYEETE-----SKLFEKIKEGYEFESP 241  
Db 191 VVGRQRYGRPVDCAWIGVIMYILSLGNPPFYEVEEDDYENHDKNLFRKILAGDYEDFSP 250

QY 242 FWDIISAKDFTCHLEKDPNRYCEKALSHPWIDGNLTALHRDIYPSVSLQIQKNFAK 301  
Db 251 YWDDISQAADLVTRLMVEYEQDQRTAEAEIAISHEWISGNAASDKNIKDGVCQAIEKNFAR 310

QY 302 SKWRQAFNAAVVHHMRKLIHMNLHSPGVREVENRPPETQAS--ETSRP-----SSPE 352  
Db 311 AKWKKAQVRVTTLMKRLR-----APEQSTAAQASASATDTATPCAAGGATAAAA 359  
QY 353 ITTTEAPVLDHVALPALTQPCQHGRRRPTAPGGRSLNCLVNGSL--HISSSLVPMHQGS 410

Db 360 SCATSAPEGDAARAASDNVAPADRSATPATDG--SATPATDGSVTPATDGSITPATDGS 417  
QY 411 L-----AAGPCGCCSCLNIGSKGKSSYCSRPSTLLKANKKQNFSEVWV---PVRASG 461  
Db 418 VTPVTDRSATP-----ATDGRATPATEESTVPTQSSAMLATKAAATPEPAMAQ 467  
QY 462 SSHCRAGQTG 471  
Db 468 DSTAPEGATG 477

RESULT 11  
US-10-217-357-4  
; Sequence 4, Application US/10217357  
; Publication No. US20030023063A1  
; GENERAL INFORMATION:  
; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Mathur, Brian  
; TITLE OF INVENTION: No. US20030023063A1el Human Kinases and Polynucleotides Encod  
; FILE REFERENCE: LEX-0254-USA  
; CURRENT APPLICATION NUMBER: US/10/217,357  
; CURRENT FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US/09/975,326  
; PRIOR FILING DATE: 2001-10-11  
; PRIOR APPLICATION NUMBER: US 60/239,821  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 765  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-217-357-4

Query Match 24.9%; Score 625.5; DB 9; Length 765;  
Best Local Similarity 34.8%; Pred. No. 4.1e-39;  
Matches 142; Conservative 82; Mismatches 143; Indels 41; Gaps 11;

QY 2 GRKEDDCSSWK-----KOTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI 54  
Db 365 GGPELDRICSPGNGNRCSESTLLEKIKYKIGVDGNFAVVKECIDRSTGKEFALKII 424

QY 55 KKSAPAF-RDSSLENIETAVLKIKHENIVTLEDIESTHYILVMQLVSGGELFDRILRG 113  
Db 425 DKACCGGHELIENEVSILRRKHPNIIIMVEEMETATELFLVWELVKGDLFAITSST 484

QY 114 VYTEKASLVITQOVLASVLYLHENGIVHRDLKPENLLYLT-PEENSKIMITDFGLSKMEQ 172  
Db 485 KYTERDGSAMVYNLANALRYLHGLSIVHRDIKPENLLVCEYPDGTTKSLKLGDFGLATVVE 544

QY 173 NGIMSTAGCTPGYVAPEVLAQKPYSAKVDCHSIVGITYILLCGYPPFYEET--ESKLPEK 230  
Db 545 -GLPYTVCGTPTYVAPEIETGYGLKYDIIWAAGVITYILLCGPPFPRSENQLQEDLFDQ 603

QY 231 IKEGYEFESFPDDIISAKDFICHLEKDPNRYCEKALSHPWIDGNLTALHRDIYPS 290  
Db 604 ILAGKLEFPAPYWDNITDSAKELISQMLQVNEARCTAGQILSHPWSDSDASQENMQAE 663

QY 291 VSLQIQKNF--AKSKWRQAFNAAVVHHMR-----KLHMNLH-----SPGVRPEVENRPP 338  
Db 664 VTGKLQKHFNALPKONSTTTTGVSVIMNTALDKQEIQCSKHODSGRPGMEP-ISPVP 722

QY 339 ETQASSTRPSPETTITTEAPVLDHVALPALTQPCQHGRRRPTAPG 386  
Db 723 -----SVEEIPVGEAVPAPTPPESPSTPH-CPPAAPGG 754

RESULT 12  
US-09-975-326-4  
; Sequence 4, Application US/09975326  
; Patent No. US20020128458A1  
; GENERAL INFORMATION:

```
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; TITLE OF INVENTION: No. US20020128458A1el Human Kinases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0254-USA
; CURRENT APPLICATION NUMBER: US/09/975,326
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: US 60/239,821
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 765
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-975-326-4

Query Match      24.9%; Score 625.5; DB 10; Length 765;
Best Local Similarity 34.8%; Pred. No. 4.1e-39;
Matches 142; Conservative 82; Mismatches 143; Indels 41; Gaps 11;

QY  2 GRKEEDDCSSWK-----KQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI 54
Db  365 GGPDLDRICISPEGVNGNRCSESTLLEKYIKGVIGDGNFAVVKECIDRSTGKEFALKII 424
QY  55 KKSAPF-RDSSLENEIAVLKKIKHENIVTLIEDIYESTTHYLVMLQVSGGELFDRILRG 113
Db  425 DKAKCCGKEHLIENEVSILRRVKHPNIIIMLVEEMETATLFLVLMELVKGDLFDALTSST 484
QY  114 VYTEKDASLVIQOVL SAVKYLHENGIVHRDLKPENLLYLT-PEENSKIMITDGLSKMEQ 172
Db  485 KYTERDGSAMVYNLANALRYLHGLSIVHRDIKPENLLVCEYDPGTSKLKGLDGLATVVE 544
QY  173 NGIMSTACGTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPPFYEBT--ESKLFEX 230
Db  545 -GPLYTVCGTPTVYVAPEIIAETGYGLKVDIWAAGVITYILLCGFPFRSNNLQEDLFDQ 603
QY  231 IKEYGYEFSPWDDISESAKDFICHLLEKDPNERVYTCESKALSHPWIDGNTALHRIYPS 290
Db  604 ILAGKLEFPAPYWDNITDSAKELISQMLQVNEARCTAGQILSHPWVSDDASQENNMQAE 663
QY  291 VSLQIQKNF--AKSKWQAFNAAVVHMR-----KLHMLNH-----SPGVRPEVENRPP 338
Db  564 VTGKLKQHFNNALPKQNSTTTGVSVMNTALDKEGQIFCSKHCQDSGRPGMEP-ISPVP 722
QY  339 ETQASRSPSPETITAPVLHDHVALPALTQLPCQHGRRPTAPGG 386
Db  723 -----SVEEIPVGEAVPAPTTPESPSTPH-CPPAAPGG 754

RESULT 13
US-09-934-406-2
; Sequence 2, Application US/09934406
; Publication No. US20020192204A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 19985, A NOVEL HUMAN SERINE/THREONINE
; FILE REFERENCE: 10448-078001
; CURRENT APPLICATION NUMBER: US/09/934,406
; PRIOR APPLICATION NUMBER: 60/226,740
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-934-406-2

Query Match      24.9%; Score 625.5; DB 9; Length 766;
Best Local Similarity 34.8%; Pred. No. 4.1e-39;
Matches 142; Conservative 82; Mismatches 143; Indels 41; Gaps 11;

QY  2 GRKEEDDCSSWK-----KQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI 54
Db  366 GGPDLDRICISPEGVNGNRCSESTLLEKYIKGVIGDGNFAVVKECIDRSTGKEFALKII 425
QY  55 KKSAPF-RDSSLENEIAVLKKIKHENIVTLIEDIYESTTHYLVMLQVSGGELFDRILRG 113
Db  426 DKAKCCGKEHLIENEVSILRRVKHPNIIIMLVEEMETATLFLVLMELVKGDLFDALTSST 485
QY  114 VYTEKDASLVIQOVL SAVKYLHENGIVHRDLKPENLLYLT-PEENSKIMITDGLSKMEQ 172
Db  486 KYTERDGSAMVYNLANALRYLHGLSIVHRDIKPENLLVCEYDPGTSKLKGLDGLATVVE 545
QY  173 NGIMSTACGTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPPFYEBT--ESKLFEX 230
Db  546 -GPLYTVCGTPTVYVAPEIIAETGYGLKVDIWAAGVITYILLCGFPFRSNNLQEDLFDQ 604
```

```
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; TITLE OF INVENTION: No. US20030023063A1el Human Kinases and Polynucleotides Encod1
; FILE REFERENCE: LEX-0254-USA
; CURRENT APPLICATION NUMBER: US/09/975,326
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US/09/975,326
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: US 60/239,821
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-217-357-2

Query Match      24.9%; Score 625.5; DB 9; Length 766;
Best Local Similarity 34.8%; Pred. No. 4.1e-39;
Matches 142; Conservative 82; Mismatches 143; Indels 41; Gaps 11;

QY  2 GRKEEDDCSSWK-----KQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI 54
Db  366 GGPDLDRICISPEGVNGNRCSESTLLEKYIKGVIGDGNFAVVKECIDRSTGKEFALKII 425
QY  55 KKSAPF-RDSSLENEIAVLKKIKHENIVTLIEDIYESTTHYLVMLQVSGGELFDRILRG 113
Db  426 DKAKCCGKEHLIENEVSILRRVKHPNIIIMLVEEMETATLFLVLMELVKGDLFDALTSST 485
QY  114 VYTEKDASLVIQOVL SAVKYLHENGIVHRDLKPENLLYLT-PEENSKIMITDGLSKMEQ 172
Db  486 KYTERDGSAMVYNLANALRYLHGLSIVHRDIKPENLLVCEYDPGTSKLKGLDGLATVVE 545
QY  173 NGIMSTACGTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPPFYEBT--ESKLFEX 230
Db  546 -GPLYTVCGTPTVYVAPEIIAETGYGLKVDIWAAGVITYILLCGFPFRSNNLQEDLFDQ 604
QY  231 IKEYGYEFSPWDDISESAKDFICHLLEKDPNERVYTCESKALSHPWIDGNTALHRIYPS 290
Db  605 ILAGKLEFPAPYWDNITDSAKELISQMLQVNEARCTAGQILSHPWVSDDASQENNMQAE 664
QY  291 VSLQIQKNF--AKSKWQAFNAAVVHMR-----KLHMLNH-----SPGVRPEVENRPP 338
Db  665 VTGKLKQHFNNALPKQNSTTTGVSVMNTALDKEGQIFCSKHCQDSGRPGMEP-ISPVP 723
QY  339 ETQASRSPSPETITAPVLHDHVALPALTQLPCQHGRRPTAPGG 386
Db  724 -----SVEEIPVGEAVPAPTTPESPSTPH-CPPAAPGG 755

RESULT 14
US-10-217-357-2
; Sequence 2, Application US/10217357
; Publication No. US20030023063A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; TITLE OF INVENTION: No. US20030023063A1el Human Kinases and Polynucleotides Encod1
; FILE REFERENCE: LEX-0254-USA
; CURRENT APPLICATION NUMBER: US/10/217,357
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US/09/975,326
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: US 60/239,821
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-217-357-2
```

QY 231 IKEYYEFESFWDIDISEAKDFICHLLEKDPNERVTCESKALSHPWIDCINTALHRDIYPS 290  
 Db 605 ILAKLEFPAPYWDNITDSAKELISQMLQVNEARCTAGQILSHFWVSDASQENMQAE 664  
 QY 291 VSLQIQKNF--AKSKWRQAFNAAVHHMR-----KLHMLH-----SPGVRPEVENRPP 338  
 Db 665 VTGKLKQHFNALPKQNSTTGTGVSIMTALDKEGQIFCSKHCQDSGRPGMEP-ISPVP 723  
 QY 339 ETQASETSRPSPEITTEAPVLDHVALPALTQPCQHGRRPTAPGG 386  
 Db 724 -----SVEEIPVGEAVPAPTPPESPTPH-CPPAAPGG 755

RESULT 15  
 US-09-975-326-2  
 ; Sequence 2, Application US/09975326  
 ; Patent No. US20020128458A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Turner, C. Alexander Jr.  
 ; APPLICANT: Mathur, Brian  
 ; TITLE OF INVENTION: No. US20020128458A1el Human Kinases and Polynucleotides Encoding  
 ; FILE REFERENCE: LEX-0254-USA  
 ; CURRENT APPLICATION NUMBER: US/09/975,326  
 ; CURRENT FILING DATE: 2001-10-11  
 ; PRIOR APPLICATION NUMBER: US 60/239,821  
 ; PRIOR FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 766  
 ; TYPE: PRT  
 ; ORGANISM: homo sapiens  
 US-09-975-326-2

Query Match 24.9%; Score 625.5; DB 10; Length 766;  
 Best Local Similarity 34.8%; Pred. No. 4.1e-39;  
 Matches 142; Conservative 82; Mismatches 143; Indels 41; Gaps 11;  
 QY 2 GRKEEDCSSWK-----KOTNIRKTFIFMEVLGSAFSEVFLVKQRLTKGLFALKCI 54  
 Db 366 GPELDRCISPEGVNGNRCSSESTLLEKYKIGKVGDNFAVVKECIDRSTGKEFALKII 425  
 QY 55 KKSAPF-RDSSLENEIAVLAKIKHENIVTLEDIVESTHYLYVMQLVSGGELFDRLIERG 113  
 Db 426 DKACCKGHELIENEVSILRRVRHPNIIIMEETATELFLVMELVKGDLFDALTSST 485  
 QY 114 VYTEKDALVIOQVLSAVKYLHENGIVHRDLKPENLLYLT-PEENSKIMITDFGLSKMEQ 172  
 Db 486 KYTERDGSAMVYNLANALRYLHGLSIVHRDIKPENLLVCEYPDGTSKSLKLGDFGLATVVE 545  
 QY 173 NGIMSTAGCTPGYVAPEVLAQPKYSKAVDCWSIGVITYILLGYPPEYBEET--ESKLFEP 230  
 Db 546 -GLYTVGCTPTYYAPEIIAETGYGLKVDIWAAGVITYILLGFPFPERSENNIQLDFDQ 604  
 QY 231 IKEYYEFESFWDIDISEAKDFICHLLEKDPNERVTCESKALSHPWIDCINTALHRDIYPS 290  
 Db 605 ILAKLEFPAPYWDNITDSAKELISQMLQVNEARCTAGQILSHFWVSDASQENMQAE 664  
 QY 291 VSLQIQKNF--AKSKWRQAFNAAVHHMR-----KLHMLH-----SPGVRPEVENRPP 338  
 Db 665 VTGKLKQHFNALPKQNSTTGTGVSIMTALDKEGQIFCSKHCQDSGRPGMEP-ISPVP 723  
 QY 339 ETQASETSRPSPEITTEAPVLDHVALPALTQPCQHGRRPTAPGG 386  
 Db 724 -----SVEEIPVGEAVPAPTPPESPTPH-CPPAAPGG 755

Search completed: March 14, 2003, 14:10:54  
 Job time : 19 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 14:17:20 ; Search time 566 Seconds  
(without alignments)  
9736.113 Million cell updates/sec

Title: US-09-960-643-1  
Perfect score: 2447  
Sequence: 1 tggagtgaggatcaagcag.....ttttctctataaaaaaaa 2447

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq.101002.*			
1:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*		
2:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*		
3:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*		
4:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*		
5:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*		
6:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*		
7:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*		
8:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*		
9:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*		
10:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*		
11:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*		
12:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*		
13:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*		
14:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*		
15:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*		
16:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*		
17:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*		
18:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*		
19:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*		
20:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*		
21:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*		
22:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*		
23:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*		
24:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2447	100.0	2447	24 AAD36140	Human calmodulin k
2	1687	68.9	1956	22 AAI60703	Human polynucleoti
3	1144	46.8	2165	22 AAI58917	Human polynucleoti
4	772	31.5	2689	22 AAS31014	Human diagnostic a
5	355	14.5	588	23 ABK43480	DNA encoding novel
6	287	11.7	362	20 AAX51838	Human secreted pro
7	274	11.2	476	22 AAS27154	cDNA encoding nove
8	274	11.2	476	23 ABK43797	DNA encoding novel
9	149	6.1	1361	22 AAH64901	Human secreted pro

10	107	4.4	373	24	ABL99884	Human secretory po
11	60	2.5	60	24	ABN35390	Human spliced tran
12	32	1.3	1512	23	AAS70157	DNA encoding novel
13	32	1.3	1774	23	AAS68767	DNA encoding novel
14	32	1.3	1774	23	AAS90919	DNA encoding novel
15	32	1.3	1774	23	AAS90919	DNA encoding novel
16	25	1.0	25	24	AAD36148	Human CAMK-X1 full
17	25	1.0	25	24	AAD36155	Human CAMK-X1 full
18	25	1.0	30	24	AAD36157	Human CAMK-X1 full
19	25	1.0	30	24	AAD36158	Human CAMK-X1 dele
20	24	1.0	31	24	AAD36156	Human CAMK-X1 full
21	23	0.9	30	24	AAD36159	Human CAMK-X1 dele
22	23	0.9	85	22	AAD36159	Human CAMK-X1 dele
23	23	0.9	296	19	AAS36908	Human colon cancer
24	23	0.9	534	21	AAV04278	Secreted protein C
25	23	0.9	3380	20	AAD00307	Arabidopsis SERK i
26	22	0.9	755	17	AAT41464	DNA encoding a FAN
27	22	0.9	1420	22	AAL13806	Human foetal liver
28	22	0.9	16759	24	ABL36297	Human cDNA sequenc
29	21	0.9	1032	22	AAS06711	Human lysosomal ac
30	21	0.9	1282	22	AAS06835	Polynucleotide seq
31	21	0.9	18183	22	AAS46835	Disease associated
32	21	0.9	18183	24	ABK31158	Tumour suppressor
33	20	0.8	352	22	ABK31158	Chemically treated
34	20	0.8	358	24	ABK8729	Signal transductio
35	20	0.8	445	22	ABK8729	Human polynucleoti
36	20	0.8	445	22	ABK8729	Arabidopsis thalia
37	20	0.8	445	22	ABK8729	Human immune/haema
38	20	0.8	445	22	ABK8729	Human immune/haema
39	20	0.8	445	22	ABK8729	Human immune/haema
40	20	0.8	1157	24	ABQ24015	Arabidopsis thalia
41	20	0.8	1157	24	ABQ24015	Oligonucleotide fo
42	20	0.8	2144	21	AAQ73212	Human 17723 recept
43	20	0.8	2334	22	AAQ74277	Rice calcium depen
44	20	0.8	2550	15	AAQ71333	Mature murine deve
45	20	0.8	3116	16	AAQ88688	Rat very low densi
			3307	19	AAV55888	Receptor protein t

ALIGNMENTS

RESULT 1  
AAD36140  
ID AAD36140 standard; DNA; 2447 BP.  
XX  
AC AAD36140;  
XX  
DT 09-AUG-2002 (first entry)  
XX  
XX Human calmodulin kinase, CAMK-X1 gene.  
DE  
DE Human; cytosolic; antisense gene therapy; screening; protein kinase;  
KW cancer; liver; colon; tumour; inflammation; arthritic synovium; CAMK-X1;  
KW calmodulin kinase; enzyme; gene; chromosome 1q32.1-32.3; ds.  
XX  
XX Homo sapiens.

XX Key Location/Qualifiers  
FH 70..1500  
FT CDS  
FT /\*tag= a  
FT /product= "Human CAMK-X1 protein"  
XX  
XX WC200224947-A2.

XX  
XX 28-MAR-2002.  
XX  
XX 20-SEP-2001; 2001WO-IB02237.  
XX  
XX 20-SEP-2000; 2000US-233999P.  
PR 02-OCT-2000; 2000US-237419P.  
PR 02-OCT-2000; 2000US-237423P.  
PR 04-OCT-2000; 2000US-238558P.  
PR 10-MAY-2001; 2001US-290555P.

XX (KINE-) KINETEK PHARM INC.  
PA (YFBR-) UNIV BRITISH COLUMBIA.  
XX  
PI Yoganathan T, Delaney AD;  
DR WPI: 2002-394145/42.  
DR P-PSDB; RAE22764.  
XX  
XX Diagnosing cancer, comprises determining the upregulation of expression  
PT of a nucleic acid sequence encoding a protein kinase or upregulation of  
PT expression of the protein kinase, in the cancer -  
XX  
PS Claim 16; Page 62-64; 87pp; English.  
XX  
CC The invention relates to a method for screening biologically active agent  
CC that modulates cancer associated protein kinase function. The invention  
CC also relates to a method for diagnosing cancer comprising determining the  
CC upregulation of expression of a nucleic acid sequence encoding a protein  
CC kinase. The method is useful for diagnosing cancer. A protein kinase is  
CC useful for screening biological agents that modulate cancer associated  
CC protein kinase function. Downregulating the activity of protein kinase is  
CC useful for inhibiting the growth of a cancer cell, e.g. liver or colon  
CC cancer. A nucleic acid encoding protein kinase is useful to screen biopsy  
CC derived tumours and inflammatory samples such as arthritic synovium, for  
CC amplified DNA in the cell or increased expression of corresponding mRNA  
CC or protein and is also useful to detect differences in expression levels  
CC such as molecular weight, amino acid and nucleotide sequences between the  
CC two cells. The present sequence is human calmodulin kinase CAMK-X1 gene  
CC located on chromosome 1q32.1-32.3.  
XX  
SQ Sequence 2447 BP; 590 A; 707 C; 604 G; 546 T; 0 other;

Query Match 100.0%; Score 2447; DB 24; Length 2447;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGAGTGGGAGCTCAAGCAGAGATTCTCCCGAGTCCCTGGCATCCTCAGAGCTTCAACT 60  
DB 1 TGGAGTGGGAGCTCAAGCAGAGATTCTCCCGAGTCCCTGGCATCCTCAGAGCTTCAACT 60  
QY 61 CTGGAGGCAATGGGTCCGAAGGAAGAAGATGACTGCTGCTGGGAAGAAACAGACACC 120  
DB 61 CTGGAGGCAATGGGTCCGAAGGAAGAAGATGACTGCTGCTGGGAAGAAACAGACACC 120  
QY 121 AACATCCGGAAAACTTCATTTTATGGAAGTGTGGGATCAGGAGCTTCTCAGAAGTT 180  
DB 121 AACATCCGGAAAACTTCATTTTATGGAAGTGTGGGATCAGGAGCTTCTCAGAAGTT 180  
QY 181 TTCCTGTGGAAGCAAGACTGACTGGAAGCTCTTCTGCTGAAGTGCATCAAGAAGTCA 240  
DB 181 TTCCTGTGGAAGCAAGACTGACTGGAAGCTCTTCTGCTGAAGTGCATCAAGAAGTCA 240  
QY 241 CCTGCTCTCCGGGACAGCAGCTGGAGAATGAGATTGCTGTGTTGAAAAGATCAAGCAT 300  
DB 241 CCTGCTCTCCGGGACAGCAGCTGGAGAATGAGATTGCTGTGTTGAAAAGATCAAGCAT 300  
QY 301 GAAACATTTGACCTTGGAGGACATCTATGAGAGCACCACCTACTACTCTGTCATG 360  
DB 301 GAAACATTTGACCTTGGAGGACATCTATGAGAGCACCACCTACTACTCTGTCATG 360  
QY 361 CAGCTTGTCTGTGGGGAGCTTTTGACCGGATCTGGAGCGGGTGTCTACACAGAG 420  
DB 361 CAGCTTGTCTGTGGGGAGCTTTTGACCGGATCTGGAGCGGGTGTCTACACAGAG 420  
QY 421 AAGATGCCAGTGTGGTATCCAGCAGGCTTTCGCGCAGTGAATATACATGAGAA 480  
DB 421 AAGATGCCAGTGTGGTATCCAGCAGGCTTTCGCGCAGTGAATATACATGAGAA 480  
QY 481 GGCATCTCCACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCTGAGAGAC 540  
DB 481 GGCATCTCCACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCTTGAAGAAC 540

QY 541 TCTAAGATCATGATCAGTACTTGGTCTGTCTCAAGATGGACAGAAATGSCATCATGTCC 600  
DB 541 TCTAAGATCATGATCAGTACTTGGTCTGTCTCAAGATGGACAGAAATGSCATCATGTCC 600  
QY 601 ACTGCTGTGGGACCCAGGCTAGCTGTGCTCCAGAAAGTGTGGCCCAAGAAACCTACAGC 660  
DB 601 ACTGCTGTGGGACCCAGGCTAGCTGTGCTCCAGAAAGTGTGGCCCAAGAAACCTACAGC 660  
QY 661 AAGGTGTGGATGCTGTGCTCATACCTACCTACATATTTGCTGTGTGGATACCC 720  
DB 661 AAGGTGTGGATGCTGTGCTCATACCTACCTACATATTTGCTGTGTGGATACCC 720  
QY 721 CCGTTCTATGAAGAAACGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAG 780  
DB 721 CCGTTCTATGAAGAAACGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAG 780  
QY 781 TTTGAGTCTCCATTCTTGGATGACATTTCTGAGTCAGCCAAAGGACTTTATTTGCCACTTG 840  
DB 781 TTTGAGTCTCCATTCTTGGATGACATTTCTGAGTCAGCCAAAGGACTTTATTTGCCACTTG 840  
QY 841 CTTGAGAGGATCCGAAAGCGCTACACCTGTGAGAGGCTTTGAGTCATCCTCGATT 900  
DB 841 CTTGAGAGGATCCGAAAGCGCTACACCTGTGAGAGGCTTTGAGTCATCCTCGATT 900  
QY 901 GACGAAACACGGCCCTCCACCGGACATCTACCCATCAGTCAGCTCCAGATCCAGAAG 960  
DB 901 GACGAAACACGGCCCTCCACCGGACATCTACCCATCAGTCAGCTCCAGATCCAGAAG 960  
QY 961 RACITTTGATAGAGCAAGTGGAGCAAGCTTCAACCCAGCAGCTGTGTGACACCATG 1020  
DB 961 RACITTTGATAGAGCAAGTGGAGCAAGCTTCAACCCAGCAGCTGTGTGACACCATG 1020  
QY 1021 AGGAAGCTTACATGAACTGCACAGCCCGGGCTGCGCCAGAGTGGAGAGACAGCCG 1080  
DB 1021 AGGAAGCTTACATGAACTGCACAGCCCGGGCTGCGCCAGAGTGGAGAGACAGCCG 1080  
QY 1081 CCTGAACTCAAGCTCAGAAACCTCTAGACCAGCTCCCTGAGATCAGCTACCGGAG 1140  
DB 1081 CCTGAACTCAAGCTCAGAAACCTCTAGACCAGCTCCCTGAGATCAGCTACCGGAG 1140  
QY 1141 GCACCTGTCTGGACACAGTGTAGCACTCCCTGCTGACCCCAATTTACCTTGCACGAT 1200  
DB 1141 GCACCTGTCTGGACACAGTGTAGCACTCCCTGCTGACCCCAATTTACCTTGCACGAT 1200  
QY 1201 GCGCGCGGCGCACTGCGCTGTGGGAGGTCCCTCAACTGCTGTGTCATGCTCCTC 1260  
DB 1201 GCGCGCGGCGCACTGCGCTGTGGGAGGTCCCTCAACTGCTGTGTCATGCTCCTC 1260  
QY 1261 CACATCAGCAGCAGCTGCTGCCATGCATCAGGGTCCCTGGCGCGCGGCGCTGTGGC 1320  
DB 1261 CACATCAGCAGCAGCTGCTGCCATGCATCAGGGTCCCTGGCGCGCGGCGCTGTGGC 1320  
QY 1321 TGCTGCTCAGCTGCCTGAACTTTGGGAGCAAGAAAGTCTCTACTCTGCTGAGCCC 1380  
DB 1321 TGCTGCTCAGCTGCCTGAACTTTGGGAGCAAGAAAGTCTCTACTCTGCTGAGCCC 1380  
QY 1381 ACATCTCTAAAAGGCAACAAAAACAGAACTTCAAGTCGGAGGTCTGTTACCGATT 1440  
DB 1381 ACATCTCTAAAAGGCAACAAAAACAGAACTTCAAGTCGGAGGTCTGTTACCGATT 1440  
QY 1441 AAAGCAGTGGCAGCTCCCTGCTGCGGAGGAGGAGTCTGCTCTCATTTATGTGA 1500  
DB 1441 AAAGCAGTGGCAGCTCCCTGCTGCGGAGGAGGAGTCTGCTCTCATTTATGTGA 1500  
QY 1501 TTCTGGAGCTGTGCTGTATGCTCAATTTTCAGAGACATATTCAACTCCTCTGCT 1560  
DB 1501 TTCTGGAGCTGTGCTGTATGCTCAATTTTCAGAGACATATTCAACTCCTCTGCT 1560  
QY 1561 CTTCCAAACCTGTGTCTATCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGCTT 1620  
DB 1561 CTTCCAAACCTGTGTCTATCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGCTT 1620  
QY 1621 AGCAGGAGCAGTTTCTGGCCAGAAAGCAGCTGTGCTGCCAGGCGGCGCCCTCATAG 1680

Db 1621 AGCAGAGAGTTTCTGGCCAGACAGACAGCTCTGCCAGCGGGCAGCCCTCATAG 1680  
QY 1681 GAGGCCAGGAGGAGCCCAAGGCGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAGC 1740  
Db 1681 GAGGCCAGGAGGAGCCCAAGGCGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAGC 1740  
QY 1741 GTGGCCACAGCTTCCAGGTTCTCCCTGACCTGCTGCTATGCCCCACACCTACGTG 1800  
Db 1741 GTGGCCACAGCTTCCAGGTTCTCCCTGACCTGCTGCTATGCCCCACACCTACGTG 1800  
QY 1801 CCGTGGCTCTGTGCAGTGTACGTAGATAGTCTTCGCCCTGGGTCTGTGCTGTTCGTGA 1860  
Db 1801 CCGTGGCTCTGTGCAGTGTACGTAGATAGTCTTCGCCCTGGGTCTGTGCTGTTCGTGA 1860  
QY 1861 AAAGCTTAATGGCTGGCCAGGCTGTGCACCTTCTCAAGCAAGCCATATGAGGATC 1920  
Db 1861 AAAGCTTAATGGCTGGCCAGGCTGTGCACCTTCTCAAGCAAGCCATATGAGGATC 1920  
QY 1921 TACCAGACTCCCACTCTGCACACTCACTCCGACCTCTCAAGCCTCCACCTCTTGGC 1980  
Db 1921 TACCAGACTCCCACTCTGCACACTCACTCCGACCTCTCAAGCCTCCACCTCTTGGC 1980  
QY 1981 CAGATTGGCTCATTAATGTCTGCTGCCCTGCCATCTGCATGAATGACAGGAGCTCCCA 2040  
Db 1981 CAGATTGGCTCATTAATGTCTGCTGCCCTGCCATCTGCATGAATGACAGGAGCTCCCA 2040  
QY 2041 TGGTGGCTGTGCTGTGAGCTTCAAGTTCTAAGCTTAACTCCAGGATGAGTCCCAAG 2100  
Db 2041 TGGTGGCTGTGCTGTGAGCTTCTCAAGTTCTAAGCTTAACTCCAGGATGAGTCCCAAG 2100  
QY 2101 TGGCTGTGAGACCCAGCAGACACTTCTGGCCCTTCTCCCTGCTCAATCTAAAGCAGT 2160  
Db 2101 TGGCTGTGAGACCCAGCAGACACTTCTGGCCCTTCTCCCTGCTCAATCTAAAGCAGT 2160  
QY 2161 GCCACACCTCCAAAGTGGATAGAAAGATTCATGAGTAAGGCTGCAAGGAATCTT 2220  
Db 2161 GCCACACCTCCAAAGTGGATAGAAAGATTCATGAGTAAGGCTGCAAGGAATCTT 2220  
QY 2221 ATCTGGCCACATGTCTCCGTGACACACCAATGAGTTAACTTGGAGTTGACTAT 2280  
Db 2221 ATCTGGCCACATGTCTCCGTGACACACCAATGAGTTAACTTGGAGTTGACTAT 2280  
QY 2281 TTAAATGTGCGCAGGATTTAAATCTGCTGTTCCTCTCTCCCTTCTCCCTGAAAGTCC 2340  
Db 2281 TTAAATGTGCGCAGGATTTAAATCTGCTGTTCCTCTCTCCCTTCTCCCTGAAAGTCC 2340  
QY 2341 AGCACACCATTTCTGCTTCCAGTTTCTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTC 2400  
Db 2341 AGCACACCATTTCTGCTTCCAGTTTCTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTC 2400  
QY 2401 AGTGTGTGCTTAATAAATGACATATTTTCTCTAAATAAAAAA 2447  
Db 2401 AGTGTGTGCTTAATAAATGACATATTTTCTCTAAATAAAAAA 2447

RESULT 2  
AAI60703  
ID AAI60703 standard; cDNA; 1956 BP.  
AC AAI60703;  
XX AAI60703;  
XX AAI60703;  
DT 22-OCT-2001 (first entry)  
DE Human polynucleotide SEQ ID NO 4692.  
KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX

OS Homo sapiens.  
XX WO200153312-A1.  
PN 26-JUL-2001.  
PD 26-JUL-2001.  
PF 26-DEC-2000; 2000WO-US34263.  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX WPI: 2001-442253/47.  
DR P-PSDB; AAM41547.  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
PT  
XX Claim 1; SEQ ID NO 4692; 10078pp; English.  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AA42213) with neurotropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 1956 BP; 474 A; 559 C; 508 G; 414 T; 1 other;  
Query Match 68.9%; Score 1687; DB 22; Length 1956;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1787; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 12 CTCAAGCAGGATTTCTCCCGAGTCCCTGGCATCTCTCAGAAGCTTCAACTCTTGAGGCAAT 71  
Db 7 CTCAAGCAGGATTTCTCCCGAGTCCCTGGCATCTCTCAGAAGCTTCAACTCTTGAGGCAAT 66  
QY 72 GGTCGAAAGGAAGAGATGACTGCGAGTCTCTGGGAAGAACACACCAATCCGGAA 131  
Db 67 GGTCGAAAGGAAGAGATGACTGCGAGTCTCTGGGAAGAACACACCAATCCGGAA 126  
QY 132 AACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTTTCTCTGGTGA 191  
Db 127 AACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTTTCTCTGGTGA 186  
QY 192 GCAAGACGACTGCGGAAGCTTTCTCTGAGTGATCATCAAGAGTCACTGCTTCG 251  
Db 187 GCAAGACGACTGCGGAAGCTTTCTCTGAGTGATCATCAAGAGTCACTGCTTCG 246  
QY 252 GGACAGCAGCTGGGAAGTCTGCTGTGTTGAAAGATCAAGCATGAAACATGTT 311  
Db 247 GGACAGCAGCTGGGAAGTCTGCTGTGTTGAAAGATCAAGCATGAAACATGTT 306

QY 312 GACCCTGGAGGACATCTATGAGAGCACCACCACCTACTACTTGGTTCATGCGAGCTGTTTC 371  
|||||  
Db 307 GACCCTGGAGGACATCTATGAGAGCACCACCACCTACTACTTGGTTCATGCGAGCTGTTTC 366  
QY 372 TGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGATGCCAG 431  
|||||  
Db 367 TGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGATGCCAG 426  
QY 432 TCTGGTGATCCAGCAGGCTTGTGGCAGTGAATACCTACATGAGATGCGATCGTCA 491  
|||||  
Db 427 TCTGGTGATCCAGCAGGCTTGTGGCAGTGAATACCTACATGAGATGCGATCGTCA 486  
QY 492 CAGAGACTTTAAAGCCGGAACCTGCTTTACCTTACCCTGAAGAGAAGCTCTAAGATCAT 551  
|||||  
Db 487 CAGAGACTTTAAAGCCGGAACCTGCTTTACCTTACCCTGNAAGAGAAGCTCTAAGATCAT 546  
QY 552 GATCACTGACTTTGGTCTCTCAAGATGGAACAGAAATGGCAATGTCATGCCGTGG 611  
|||||  
Db 547 GATCACTGACTTTGGTCTCTCAAGATGGAACAGAAATGGCAATGTCATGCCGTGG 606  
QY 612 GACCCAGGCTACGTGGCTCCAGAGTGTGGCCAGAAACCTACAGCAAGGCTGTGA 671  
|||||  
Db 607 GACCCAGGCTACGTGGCTCCAGAGTGTGGCCAGAAACCTACAGCAAGGCTGTGA 666  
QY 672 TTGCTGGTCCATCGGGCTCATCACCTACATATTGCTGTGGATACCCCGCTTCTATGA 731  
|||||  
Db 667 TTGCTGGTCCATCGGGCTCATCACCTACATATTGCTGTGGATACCCCGCTTCTATGA 726  
QY 732 AGAAACGGAGTCTAAGCTTTTCAGAGATCAAGAGGGCTACTATGATTTGATGCTCC 791  
|||||  
Db 727 AGAAACGGAGTCTAAGCTTTTCAGAGATCAAGAGGGCTACTATGATTTGATGCTCC 786  
QY 792 ATTCTCGGATGACATTTCTGAGTCAAGCAAGACCTTTATTGCCACTTGCCTGGAAGA 851  
|||||  
Db 787 ATTCTCGGATGACATTTCTGAGTCAAGCAAGACCTTTATTGCCACTTGCCTGGAAGA 846  
QY 852 TCGAAGCAGCGGTACACCTGTGAGAAGGCTTGGTGCATCCCTGGATTGACGGAAACAC 911  
|||||  
Db 847 TCGAAGCAGCGGTACACCTGTGAGAAGGCTTGGTGCATCCCTGGATTGACGGAAACAC 906  
QY 912 GGCCCTCCACGGGACATCTACCATCATGTCAGCTCAGCTCAGATCCAGAGAACTTTGCTAA 971  
|||||  
Db 907 AGCCCTCCACGGGACATCTACCATCATGTCAGCTCAGCTCAGATCCAGATPCCAGAAACTTTGCTAA 966  
QY 972 GAGCAAGTGGAGGAGCCCTTCAACGACGACGCTGTGGTGCACCATCATGAGGAAGCTACA 1031  
|||||  
Db 967 GAGCAAGTGGAGGAGCCCTTCAACGACGACGCTGTGGTGCACCATCATGAGGAAGCTACA 1026  
QY 1032 CATGAACCTGCACAGCCCGGGGCTCGGCCAGAGGTGGAGAACAGCGCCCTGAAACTCA 1091  
|||||  
Db 1027 CATGAACCTGCACAGCCCGGGGCTCGGCCAGAGGTGGAGAACAGCGCCCTGAAACTCA 1086  
QY 1092 AGCCTCAGAAACCTCTAGACCCAGCTCCCTCAGATCACCATCACCGAGGACCTGCTCT 1151  
|||||  
Db 1087 AGCCTCAGAAACCTCTAGACCCAGCTCCCTCAGATCACCATCACCGAGGACCTGCTCT 1146  
QY 1152 GGACCAAGTGTAGCACTCCCTGCCCTGACCAATTTACCCCTGCCAGTGGCCGCGGCC 1211  
|||||  
Db 1147 GGACCAAGTGTAGCACTCCCTGCCCTGACCAATTTACCCCTGCCAGTGGCCGCGGCC 1206  
QY 1212 CACTGCCCTGTGGCAGGTCCCTCAACTGCTTCAATGGCTCCCTCCACATCAGCAG 1271  
|||||  
Db 1207 CACTGCCCTGTGGCAGGTCCCTCAACTGCTTCAATGGCTCCCTCCACATCAGCAG 1266  
QY 1272 CAGCCTGGTGGCCATCATCAGGGTCCCTGGCCCGGGGCTGTGGCTGCTGCTCCAG 1331  
|||||  
Db 1267 CAGCCTGGTGGCCATCATCAGGGTCCCTGGCCCGGGGCTGTGGCTGCTGCTCCAG 1326  
QY 1332 CTGCCCTGAACATTTGGGAGCAAGAAAGTCTCTACTGCTGTAGGCCACACTCCTCAA 1391  
|||||  
Db 1327 CTGCCCTGAACATTTGGGAGCAAGAAAGTCTCTACTGCTGTAGGCCACACTCCTCAA 1386

QY 1392 AAAGGCCAACAAAAACAGAACTTCAAGTCGAGGTCATGGTACCAGTTAAAGCCAGTGG 1451  
|||||  
Db 1387 AAAGGCCAACAAAAACAGAACTTCAAGTCGAGGTCATGGTACCAGTTAAAGCCAGTGG 1446  
QY 1452 CAGTCCCACTGGCGGCGAGGAGACTGGAGTCTGTCTCATTTATGTGATTCCTGGAGCC 1511  
|||||  
Db 1447 CAGTCCCACTGGCGGCGAGGAGACTGGAGTCTGTCTCATTTATGTGATTCCTGGAGCC 1506  
QY 1512 TGTGCTTATGTCACTGCAATTTTCAGGAGACATATCAACTCTCTGCTCTTCCAAACCT 1571  
|||||  
Db 1507 TGTGCTTATGTCACTGCAATTTTCAGGAGACATATCAACTCTCTGCTCTTCCAAACCT 1566  
QY 1572 GGTGCTTATCCGCGAGAGGAGGAGGAGCAAGTGGAGCAGGGCTTAGCAGGAGCAG 1631  
|||||  
Db 1567 GGTGCTTATCCGCGAGAGGAGGAGGAGGAGCAAGTGGAGCAGGGCTTAGCAGGAGCAG 1626  
QY 1632 TTTCTGCCAGAGACACCGCTGTGCCAGCGGGGAGGCCCTCATAGGAGGCCCAGGA 1691  
|||||  
Db 1627 TTTCTGCCAGAGACACCGCTGTGCCAGCGGGGAGGCCCTCATAGGAGGCCCAGGA 1686  
QY 1692 GGGAGCCCCAAGCGGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACCA 1751  
|||||  
Db 1687 GGGAGCCCCAAGCGGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACCA 1746  
QY 1752 GCTTCCAGGTCTCCCTGACCTGCTGCTATGCCCCACACCCCTACGTG 1800  
|||||  
Db 1747 GCTTCCAGGTCTCCCTGACCTGCTGCTATGCCCCACACCCCTACGTG 1795

## RESULT 3

AAI58917  
ID AAI58917 standard; cdna; 2165 BP.  
XX AC AAI58917;  
XX AC AAI58917;  
DF 22-OCT-2001 (first entry)  
XX XX  
DE Human polynucleotide SEQ ID NO 1120.  
XX XX  
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX XX  
OS Homo sapiens.  
XX XX  
PN WO200153312-A1.  
XX XX  
PD 26-JUL-2001.  
XX XX  
PF 26-DEC-2000; 2000WO-US34263.  
XX XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX XX  
PA (HYSE-) HYSEQ INC.  
XX XX  
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX XX  
DR WPI: 2001-442253/47.  
DR P-PSDB; AAM39761.  
XX XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -  
PS Claim 1; SEQ ID NO 1120; 10078pp; English.  
XX  
XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AA138642-AA142113) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities, such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 2165 BP; 533 A; 613 C; 563 G; 455 T; 1 other;  
Query Match 46.8%; Score 1144; DB 22; Length 2165;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1294; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 504 GCCGAAAACCTGCTTACCTTACCCCTGAAGAGAACTCTAAGATCATGACTGACTT 563  
DB 708 GCCGAAAACCTGCTTACCTTACCCCTGAAGAGAACTCTAAGATCATGACTGACTT 767  
QY 564 TGCTCTGTCACAGATGACAGAAAGGCATCATGTCCTGCTGGAGCCCGAGCTA 623  
DB 768 TGCTCTGTCACAGATGACAGAAAGGCATCATGTCCTGCTGGAGCCCGAGGATA 827  
QY 624 CGTGGCTCCAGAAAGTCTGGCCAGAAACCTTACAGAAAGCTGTGGATTCCTGCTCAT 683  
DB 828 CGTGGCTCCAGAAAGTCTGGCCAGAAACCTTACAGAAAGCTGTGGATTCCTGCTCAT 887  
QY 684 CGCGCTACACCTACATATGCTCTGTGATACCCCGTCTCTATGAAGAAAGGATC 743  
DB 888 CGCGCTACACCTACATATGCTCTGTGATACCCCGTCTCTATGAAGAAAGGATC 947  
QY 744 TAAGCTTTTCGAGAAGTCAAGGAGGCTACTATGAGTTTGAGTCTCCATTTCTGGGATGA 803  
DB 948 TAAGCTTTTCGAGAAGTCAAGGAGGCTACTATGAGTTTGAGTCTCCATTTCTGGGATGA 1007  
QY 804 CATTTCTGAGTCAGCAAGGACTTTATTTGCCACTTGTGAGAGGATCCGAGGAGCG 863  
DB 1008 CATTTCTGAGTCAGCAAGGACTTTATTTGCCACTTGTGAGAGGATCCGAGGAGCG 1067  
QY 864 GTACACCTGTGAGAAGGCTTGTAGTTCATCCCTGGATTGACGGAACACGGCCCTCCACCG 923  
DB 1068 GTACACCTGTGAGAAGGCTTGTAGTTCATCCCTGGATTGACGGAACACAGCCCTCCACCG 1127  
QY 924 GGACATCTACCCATCAGTCAGCTCCAGATCCAGAGAACTTTGCTTAAGAGCAAGTGAG 983  
DB 1128 GGACATCTACCCATCAGTCAGCTCCAGATCCAGAGAACTTTGCTTAAGAGCAAGTGAG 1187  
QY 984 GCAAGCTTCAACGACAGCAGCTGTGTGACACCATGAGGAGCTACACATGAACCTGCA 1043  
DB 1188 GCAAGCTTCAACGACAGCAGCTGTGTGACACCATGAGGAGCTACACATGAACCTGCA 1247  
QY 1044 CAGCCCGGGGCTCCGCCAGAGGTGGAGAACAGGGCCCTGAACTCAAGCCTCAGAAAC 1103  
DB 1248 CAGCCCGGGGCTCCGCCAGAGGTGGAGAACAGGGCCCTGAACTCAAGCCTCAGAAAC 1307  
QY 1104 CTCTAGACCCAGCTCCCTGAGATCACCATCACCAGGACCTGCTCGGACCAAGTGT 1163  
DB 1308 CTCTAGACCCAGCTCCCTGAGATCACCATCACCAGGACCTGCTCGGACCAAGTGT 1367  
QY 1164 AGCACTCCCTGCCCCTGACCAAAATACCTGTCACAGATGCCCGCGCCCACTGCCCTGG 1223

DB 1368 AGCACTCCCTGCCCTGACCCAAATTACCTGCGACGATGGCGCGCGCCCACTGCCCTTGG 1427  
QY 1224 TGGCAGGTCCTCAACTGCTGCTCAATGGCTCCCTCCACATCAGCAGCAGCTGGTGGCC 1383  
DB 1428 TGGCAGGTCCTCAACTGCTGCTCAATGGCTCCCTCCACATCAGCAGCAGCTGGTGGCC 1487  
QY 1284 CATGATCAGGGGTCCCTGGCCGCGCGGCTGTGGCTGTGCTGCCAGCTCCCTGAACAT 1343  
DB 1488 CATGATCAGGGGTCCCTGGCCGCGCGGCTGTGGCTGTGCTGCCAGCTCCCTGAACAT 1547  
QY 1344 TGGGACCAAGAAAGTCTCTACTGCTCTGAGCCACACTCTCTCAAAAAGGCCAACAA 1403  
DB 1548 TGGGACCAAGAAAGTCTCTACTGCTCTGAGCCACACTCTCTCAAAAAGGCCAACAA 1607  
QY 1404 AAACAGAACTTCAAGTCGGAGGTATGATACCACTTAAAGCCAGTGGCAGCTCCCACTG 1463  
DB 1608 AAACAGAACTTCAAGTCGGAGGTATGATACCACTTAAAGCCAGTGGCAGCTCCCACTG 1667  
QY 1464 CCGGCGAGGCGAGACTGGAGTCTGCTCATATTGATGATTCCTGGAGCCTGTGCCTATGTC 1523  
DB 1668 CCGGCGAGGCGAGACTGGAGTCTGCTCATATTGATGATTCCTGGAGCCTGTGCCTATGTC 1727  
QY 1524 ACTGCAATTTTCAGGAGACATATTCAACTCTCTCTCTTCCAAACCTGGTGTCTATCCG 1583  
DB 1728 ACTGCAATTTTCAGGAGACATATTCAACTCTCTCTCTTCCAAACCTGGTGTCTATCCG 1787  
QY 1584 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1643  
DB 1788 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1847  
QY 1644 AGCACCAGGCTGCTGCGAGCGGCGAGCCCTCATAGGAGGCCAGGAGGAGGAGGAGGAGGAG 1703  
DB 1848 AGCACCAGGCTGCTGCGAGCGGCGAGCCCTCATAGGAGGCCAGGAGGAGGAGGAGGAGGAG 1907  
QY 1704 GCGTGAAGGCTGTTGTAAGCTGTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1763  
DB 1908 GCGTGAAGGCTGTTGTAAGCTGTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1967  
QY 1764 CCGTGAAGGCTGTTGTAAGCTGTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800  
DB 1968 CCGTGAAGGCTGTTGTAAGCTGTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2004  
RESULT 4  
AAS31014  
ID AAS31014 standard; cdna; 2689 BP.  
XX  
AC AAS31014;  
XX  
DT 04-DEC-2001 (first entry)  
XX  
DE Human diagnostic and therapeutic polynucleotide (DITHP) #29.  
XX  
KW Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;  
KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;  
KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;  
XX respiratory disorder; ss.  
OS Homo sapiens.  
XX  
PN WO200162927-A2.  
XX  
PD 30-AUG-2001.  
XX  
XX 21-FEB-2001; 2001WO-US06059.  
XX  
XX 24-FEB-2000; 2000US-0184693.  
PR 24-FEB-2000; 2000US-0184697.  
PR 24-FEB-2000; 2000US-0184698.  
PR 24-FEB-2000; 2000US-0184768.  
PR 24-FEB-2000; 2000US-0184769.  
PR 24-FEB-2000; 2000US-0184770.  
PR 24-FEB-2000; 2000US-0184771.

PR 24-FEB-2000; 2000US-0184772.  
PR 24-FEB-2000; 2000US-0184773.  
PR 24-FEB-2000; 2000US-0184774.  
PR 24-FEB-2000; 2000US-0184776.  
PR 24-FEB-2000; 2000US-0184777.  
PR 24-FEB-2000; 2000US-0184797.  
PR 24-FEB-2000; 2000US-0184813.  
PR 24-FEB-2000; 2000US-0184837.  
PR 24-FEB-2000; 2000US-0184841.  
PR 24-FEB-2000; 2000US-0185213.  
PR 24-FEB-2000; 2000US-0185216.  
PR 12-MAY-2000; 2000US-0203785.  
PR 15-MAY-2000; 2000US-0204226.  
PR 16-MAY-2000; 2000US-0204525.  
PR 16-MAY-2000; 2000US-0204821.  
PR 16-MAY-2000; 2000US-0204908.  
PR 16-MAY-2000; 2000US-0205232.  
PR 17-MAY-2000; 2000US-0204815.  
PR 17-MAY-2000; 2000US-0204863.  
PR 17-MAY-2000; 2000US-0205221.  
PR 17-MAY-2000; 2000US-0205285.  
PR 17-MAY-2000; 2000US-0205286.  
PR 17-MAY-2000; 2000US-0205287.  
PR 17-MAY-2000; 2000US-0205323.  
PR 17-MAY-2000; 2000US-0205324.  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;  
PI Chen A, D'Sa SA, Anshay S, Dahl CR, Dam TC, Daniels SE;  
PI Dufour GE, Flores V, Fong WT, Greenawalt LB, Hillman JL, Jones AL;  
PI Liu TF, Roseberry AM, Russo BH, Russo FD, Stockdreher TK, Daffo A;  
PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;  
PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;  
XX  
XX WPI: 2001-502867/55.  
DR P-PSDB; AAU19443.  
XX  
XX Polynucleotides encoding diagnostic and therapeutic proteins, e.g.  
PT enzymes, hormones and receptors, useful in diagnostics and therapeutics  
PT -  
XX  
XX Claim 1: Page 310-311; 522pp; English.

CC The invention relates to polynucleotides (I) encoding diagnostic and  
CC therapeutic (DITHP) polypeptides (II), which include e.g. enzymes,  
CC and proteins involved in growth and development and receptors. (I) and  
CC (II) may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate DITHP expression. For example, (I) and  
CC (II) may be used to treat disorders associated with decreased polypeptide  
CC expression by rectifying mutations or deletions in a patient's genome,  
CC that affect the activity of the DITHPs, by expressing inactive proteins  
CC or supplementing the patient's own production of them. (I) and (II)  
CC may be used to treat diseases, for example, cell proliferative disorder,  
CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,  
CC leukaemia, autoimmune disorders, and respiratory disorders. Additionally,  
CC (I) may be used to produce the DITHPs, by inserting the nucleic acids  
CC into a host cell and culturing the cell to express the protein. (I) and  
CC its complementary sequences may also be used as DNA probes in diagnostic  
CC assays to detect and quantitate the presence of similar nucleic acids in  
CC samples, and therefore which patients may be in need of restorative  
CC therapy. (II) may also be used as antigens in the production of  
CC antibodies against DITHPs and in assays to identify modulators of DITHP  
CC expression and activity. The anti-DITHP antibodies and antagonists may  
CC also be used to down regulate expression and activity. The anti-DITHP  
CC antibodies may also be used as diagnostic agents for detecting the  
CC presence of DITHPs in samples (e.g. by enzyme linked immunosorbent  
CC assay (ELISA)). AAS30986-AAS31196 represent human diagnostic and  
CC therapeutic (DITHP) polynucleotides of the invention.  
XX  
XX Sequence 2689 BP; 642 A; 748 C; 678 G; 621 T; 0 other;

Query Match 31.5%; Score 772; DB 22; Length 2689;

Best Local Similarity 99.9%; Pred. No. 0;  
Matches 822; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 39 GGCACTCCTCAGAGCTTCAACTCTGGAGGCAATGGGTCGAAAGAGACATGACTGCAG 98  
DB 206 GGCATCCTCAGAGCTTCAACTCTGGAGGCAATGGGTCGAAAGAGACATGACTGCAG 265  
QY 99 TTCCTGGAAGAAACAGACACCACACATCCCGAAACCTTCATTTTATGAACTGCTGGG 158  
DB 266 TTCCTGGAAGAAACAGACACCACACATCCCGAAACCTTCATTTTATGAACTGCTGGG 325  
QY 159 ATCAGGAGCTTCTCAGAGTCTTTCCTGGTGAAGCAAGACTGACTGGGAGCTCTTTGC 218  
DB 326 ATCAGGAGCTTCTCAGAGTCTTTCCTGGTGAAGCAAGACTGACTGGGAGCTCTTTGC 385  
QY 219 TCTGAAGTGCATCAAGAAGTCACTGCTTCCGGGACAGCAGCTGGAGAAATGAGATTGC 278  
DB 386 TCTGAAGTGCATCAAGAAGTCACTGCTTCCGGGACAGCAGCTGGAGAAATGAGATTGC 445  
QY 279 TGTGTTGAAAAGATCAAGCATGAAACATTTGTACCTGGAGACATCTATGAGAGCAC 338  
DB 446 TGTGTTGAAAAGATCAAGCATGAAACATTTGTACCTGGAGACATCTATGAGAGCAC 505  
QY 339 CACCCACTACTACCTGCTCATGAGCTGTTCTGTTGGGGAGCTCTTTGACCGGATCCT 398  
DB 506 CACCCACTACTACCTGCTCATGAGCTGTTCTGTTGGGGAGCTCTTTGACCGGATCCT 565  
QY 399 GGACGGGGTGTCTACACAGAGAGGATGCCAGTCTGTGTCATCCAGCAGCTCTTGTCCGC 458  
DB 566 GGACGGGGTGTCTACACAGAGAGGATGCCAGTCTGTGTCATCCAGCAGCTCTTGTCCGC 625  
QY 459 AGTGAATACCTACATGAGAAATGGCATCGTCCACAGAGACTTAAAGCCCGAAACCTGCT 518  
DB 626 AGTGAATACCTACATGAGAAATGGCATCGTCCACAGAGACTTAAAGCCCGAAACCTGCT 685  
QY 519 TTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCACTGACTTTGCTGTGTCACAGAT 578  
DB 686 TTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCACTGACTTTGCTGTGTCACAGAT 745  
QY 579 GGAACAGAAATGGCATATGTCCTACTGCTGAGGACCCAGGCTAGCTGGCTCCAGAAAT 638  
DB 746 GGAACAGAAATGGCATATGTCCTACTGCTGAGGACCCAGGCTAGCTGGCTCCAGAAAT 805  
QY 639 GCTGGCCCAAGAAACCTTACAGCAAGGCTGTGATTTGCTGCTCCATCGGCTCATACCTA 698  
DB 806 GCTGGCCCAAGAAACCTTACAGCAAGGCTGTGATTTGCTGCTCCATCGGCTCATACCTA 865  
QY 699 CATATTGCTCTGTGATACCCCGCTTCTATGAAGAAACGAGTCTAAGCTTTTCGAGAA 758  
DB 866 CATATTGCTCTGTGATACCCCGCTTCTATGAAGAAACGAGTCTAAGCTTTTCGAGAA 925  
QY 759 GATCAAGAGGGCTTACTATGATTTGAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGC 818  
DB 926 GATCAAGAGGGCTTACTATGATTTGAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGC 985  
QY 819 CAAGGACTTTATTTGCCACTTGTCTGAGAAGGATCCGAACGAG 861  
DB 986 CAAGGACTTTATTTGCCACTTGTCTGAGAAGGATCCGAACGAG 1028  
RESULT 5  
ABK43480  
ID ABK43480 standard; cDNA; 588 BP.  
XX  
AC ABK43480;  
XX  
DT 05-JUN-2002 (first entry)  
XX  
DE DNA encoding novel central nervous system protein #60.  
XX  
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;  
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;

KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
KW adenocarcinoma; reproductive system disorder; testicular feminisation;  
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;  
KW respiratory disorder; renal disorder; kidney failure; blood disorder;  
KW myocardial infarction; wound healing; cell proliferation; skin aging;  
KW food additive; food preservative; gene therapy; gene; ss.  
XX  
OS Homo sapiens.  
PN WO200155318-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01332.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 08-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234597.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.

PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 PI WPI; 2001-581633/65.  
 DR P-PSDB; AAU87150.  
 XX  
 XX New isolated nucleic acid encoding a protein for diagnosing,  
 PT preventing, treating or ameliorating medical conditions and used as  
 PT food additives or preservatives -  
 XX  
 XX Claim 1; SEQ ID No 70; 837pp; English.

XX The invention describes an isolated nucleic acid molecule (I) encoding a  
 CC novel central nervous system protein. (I) and polypeptides (III) encoded  
 CC by (I), are used to treat a medical conditions and in diagnosis of a  
 CC pathological condition. Disorders which are diagnosed or treated include  
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and  
 CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses  
 CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders  
 CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,  
 CC adenocarcinomas and irritable bowel syndrome, reproductive system  
 CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes  
 CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.  
 CC leukaemia, disorders involving neovascularisation e.g. malignancies,  
 CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.  
 CC acute kidney failure and blood related disorders e.g. myocardial  
 CC infarction. The polypeptides can also be used to aid wound healing and  
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to  
 CC maintain organs before transplantation, for supporting cell culture of  
 CC primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities, fat content, lipid, protein,  
 Query Match 14.5%; Score 355; DB 23; Length 588;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-162;  
 Matches 355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGGAGTGGAGCTCAAGCAGGATTCTTCCGAGTCCCTGGCATCTCAGAGCTTCAACT 60  
 Db 145 TGGAGTGGAGCTCAAGCAGGATTCTTCCGAGTCCCTGGCATCTCAGAGCTTCAACT 204  
 Qy 61 CTGAGGCAATGGCTGAAAGGAGAGAGTACTGCTGCTTCTGGAGAAACAGACACC 120  
 Db 205 CTGAGGCAATGGCTGAAAGGAGAGAGTACTGCTGCTTCTGGAGAAACAGACACC 264  
 Qy 121 AACATCCGGAACCTTCATTTTATGGAAGTGGTGGATCAGAGCTTCTCAGAAGTT 180  
 Db 265 AACATCCGGAACCTTCATTTTATGGAAGTGGTGGATCAGAGCTTCTCAGAAGTT 324  
 Qy 181 TTCTGTTGGAAGCAAGAGTACTGCTGGAAGTCTTTCTGCTGAGTGCATCAAGAAGTCA 240  
 Db 325 TTCTGTTGGAAGCAAGAGTACTGCTGGAAGTCTTTCTGCTGAGTGCATCAAGAAGTCA 384  
 Qy 241 CCTGCTTCCGGGACAGACGCTGGAGAGTATGAGATGCTGTGTGAAAGAGATCAAGCAT 300  
 Db 385 CCTGCTTCCGGGACAGACGCTGGAGAGTATGAGATGCTGTGTGAAAGAGATCAAGCAT 444  
 Qy 301 GAAACATTTGACCTTGGAGGACATCTATGAGAGCACCACCTACTACTCTGG 355  
 Db 445 GAAACATTTGACCTTGGAGGACATCTATGAGAGCACCACCTACTACTCTGG 499

RESULT 6  
 AAX51838  
 ID AAX51838 standard; DNA; 362 BP.  
 XX  
 AC AAX51838;  
 XX  
 DT 22-JUN-1999 (first entry)  
 XX  
 DE Human secreted protein 5' EST SEQ ID NO: 52.  
 XX  
 KW Human; secreted protein; EST; expressed sequence tag; diagnosis;  
 KW forensic; gene therapy; chromosome mapping; signal peptide;  
 KW upstream regulatory sequence; cytokine activity; cell proliferation;  
 KW differentiation; haematopoiesis regulation; tissue growth regulation;  
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;  
 KW thrombolytic; anti-inflammatory; tumour inhibition; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9906552-A2.  
 XX  
 PD 11-FEB-1999.  
 XX  
 PF 31-JUL-1998; 98WO-IB01236.  
 XX  
 PR 01-AUG-1997; 97US-0905223.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Duclert A, Dumas Milne Edwards J, Lacroix B;  
 XX WPI; 1999-153782/13.  
 DR P-PSDB; AAY13038.  
 XX  
 XX New isolated brain-derived nucleic acids - used to develop products  
 PT which may have cytokine, immune, regulatory, haematopoiesis  
 PT regulating, anti-inflammatory or tumour inhibition activity  
 XX  
 PS Claim 1; Page 212; 577pp; English.

AAX51787 to AAX52019 represent 5' expressed sequence tags (ESTs) for  
 human secreted proteins, and encode the proteins given in AAY12987 to  
 AAY13219, respectively. The proteins given represent the signal peptide  
 and an N-terminal fragment of a secreted protein. The nucleic acid  
 sequences can be used for producing secreted human gene products. They  
 can also be used to develop products for diagnosis and therapy. The  
 proteins obtained may have cytokine activity, cell  
 proliferation/differentiation activity, haematopoiesis regulating  
 activity, tissue growth regulating activity, reproductive hormone  
 regulating activity, chemotactic/chemokinetic activity, haemostatic and  
 thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 activity, tumour inhibition activity or other activities. The products  
 can be used in forensic, gene therapy and chromosome mapping procedures.  
 The sequences can also be used for obtaining corresponding promoter  
 sequences. The nucleic acids encoding the signal peptide can be used for  
 directing extracellular secretion of a polypeptide or the insertion of a  
 polypeptide into a membrane, or importing a polypeptide into a cell.

XX Sequence 362 BP; 107 A; 79 C; 99 G; 76 T; 1 other;  
 Query Match 11.7%; Score 287; DB 20; Length 362;  
 Best Local Similarity 99.7%; Pred. No. 3.2e-129;  
 Matches 337; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 TGGAGTGGAGCTCAAGCAGGATTCTTCCGAGTCCCTGGCATCTCAGAGCTTCAACT 60  
 Db 18 TGGAGTGGAGCTCAAGCAGGATTCTTCCGAGTCCCTGGCATCTCAGAGCTTCAACT 77  
 Qy 61 CTGAGGCAATGGCTGAAAGGAGAGTACTGCTGCTTCTGGAGAAACAGACACC 120  
 Db 78 CTGAGGCAATGGCTGAAAGGAGAGTACTGCTGCTTCTGGAGAAACAGACACC 137  
 Qy 121 AACATCCGGAACCTTCATTTTATGGAAGTGGTGGATCAGAGCTTCTCAGAAGTT 180

```

|||||
Db 138 AACATCGGAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTT 197
QY 181 TTCCTGCTGAAGCAAGACTGACTGGAGCTCTTTGCTCTGAAGTGCATCAAGACTCA 240
Db 198 TTCCTGCTGAAGCAAGACTGACTGGAGCTCTTTGCTCTGAAGTGCATCAAGACTCA 257
QY 241 CCTGCCCTCCGGACAGCAGCCTGGAGATGAGATTGCTGTGTTGAAAAAGATCAAGCAT 300
Db 258 CTGCTCCCGGACAGCAGCCTGGAGATGAGATTGCTGTGTTGAAAAAGATCAAGCAT 317
QY 301 GAAAAATTTGACCTCGGAGACATCTATGAGAGCAC 338
Db 318 GAAAAATTTGACCTCGGAGACATCTATGAGAGCAC 355

RESULT 7
AAS27154
ID AAS27154 standard; cDNA; 476 bp.
AC AAS27154;
XX
DT 07-NOV-2001 (first entry)
XX
DE cDNA encoding novel signal transduction pathway protein, Seq ID 189.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200154733-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01312.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216680.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0228868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246533.
PR 08-NOV-2000; 2000US-0246609.

```

PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-465460/50.  
DR P-PSDB; AAU17237.  
XX  
XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
PT prognosing disorders related to the proteins, including cancers, immune  
PT disorders and neuronal disorders -  
XX  
XX Claim 1; SEQ ID NO 189; 880pp; English.  
XX  
XX The invention relates to novel isolated polypeptides (I), and  
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
CC diagnosing, preventing and treating diseases including immune system  
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
CC transplant rejections and graft versus host disease, infectious diseases  
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
CC other blood-related disorders (sickle cell anaemia), myeloproliferative  
CC disorders, primary haematopoietic disorders, hyperproliferative  
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
CC abnormalities (Down syndrome) ischaemic injury (e.g. stroke), renal  
CC disorders (e.g. glomerulonephritis), cardiovascular disorders  
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.  
CC Addison's disease), reproductive system disorders, gastrointestinal  
CC disorder (inflammatory disorders), liver disorders (cirrhosis),  
CC as stimulators of B-cell responsiveness to pathogens, activators of  
CC T-cells, to induce higher affinity antibodies, and as a means to induce  
CC tumour proliferation in pathologies e.g. acquired immune deficiency  
CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction  
CC pathway protein coding sequences and PCR primers of the invention.  
XX

Query Match 11.2%; Score 274; DB 22; Length 476;  
Best Local Similarity 100.0%; Pred. No. 6.9e-123;  
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TGGAGTGGGAGCTCAAGCAGGATTCTTCGGAGTCCCTGGCATCCTCAGAAGCTTCAACT 60  
|||||  
Db 145 TGGAGTGGGAGCTCAAGCAGGATTCTTCGGAGTCCCTGGCATCCTCAGAAGCTTCAACT 204  
|||||  
QY 61 CTGAGGCAATGGTGGAAAGGAAGATGACTGCAGTTCCTGGGAAGAAACAGACACC 120  
|||||  
Db 205 CTGAGGCAATGGTGGAAAGGAAGATGACTGCAGTTCCTGGGAAGAAACAGACACC 264  
|||||  
QY 121 AACATCCGGAAAAACCTTCATTTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTT 180  
|||||  
Db 265 AACATCCGGAAAAACCTTCATTTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTT 324  
|||||  
QY 181 TTCTGTGTGAAGCAAGACACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAGAAGTCA 240  
|||||  
Db 325 TTCTGTGTGAAGCAAGACACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAGAAGTCA 384  
|||||  
QY 241 CCTGCCTTCCGGGACAGCAGCTGGGAGATGAGA 274  
|||||  
Db 385 CCTGCCTTCCGGGACAGCAGCTGGGAGATGAGA 418  
|||||  
  
RESULT 8  
ABK43797  
ID ABK43797 standard; cDNA; 476 BP.  
XX  
XX AC ABK43797;  
XX  
XX DT 05-JUN-2002 (first entry)  
XX  
XX DE DNA encoding novel central nervous system protein #377.  
XX  
XX KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
KW hyperproliferative disorder; neoplasm, cardiovascular disorder;  
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
KW adenocarcinoma; reproductive system disorder; testicular feminisation;  
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;  
KW respiratory disorder; renal disorder; kidney failure; blood disorder;  
KW myocardial infarction; wound healing; cell proliferation; skin aging;  
KW food additive; food preservative; gene therapy; gene; ss.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200155318-A2.  
XX  
XX PD 02-AUG-2001.  
XX  
XX PF 17-JAN-2001; 2001WO-US01332.  
XX  
XX PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225216.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0233423.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-581633/65.  
XX P-PSDB; AAU87467.  
XX  
XX New isolated nucleic acid encoding a protein for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives -  
XX  
PS Claim 1; SEQ ID No 387; 837pp; English.  
XX  
XX The invention describes an isolated nucleic acid molecule (I) encoding a novel central nervous system protein. (I) and polypeptides (III) encoded by (I), are used to treat a medical conditions and in diagnosis of a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease and amyotrophic lateral sclerosis, infections caused by bacteria, viruses e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders e.g. corneal infection, gastrointestinal disorders e.g. dysphagia, adenocarcinomas and irritable bowel syndrome, reproductive system disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes and pituitary dwarfism, cancers and disorders at the cellular level e.g. leukaemia, disorders involving neovascularisation e.g. malignancies, CC

CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.  
CC acute kidney failure and blood related disorders e.g. myocardial  
CC infarction. The polypeptides can also be used to aid wound healing and  
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to  
CC maintain organs before transplantation, for supporting cell culture of  
CC primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities, fat content, lipid, protein,  
  
Query Match 11.2%; Score 274; DB 23; Length 476;  
Best Local Similarity 100.0%; Pred. No. 6.9e-133;  
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 TGGAGTGGGAGCTCAAGCAGGATTCTTCCGAGTCCCTGGCATCTCTCAGAAGCTTCAACT 60  
Db 145 TGGAGTGGGAGCTCAAGCAGGATTCTTCCGAGTCCCTGGCATCTCTCAGAAGCTTCAACT 204  
  
Qy 61 CTGGAGGCAATGGTGCAGGAAGATGATGACTGCAGTCTCTGGAAGAACAGACCACC 120  
Db 205 CTGGAGGCAATGGTGCAGGAAGATGATGACTGCAGTCTCTGGAAGAACAGACCACC 264  
  
Qy 121 AACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAGTT 180  
Db 265 AACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAGTT 324  
  
Qy 181 TTCTCTGGTGAAGCAAAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA 240  
Db 325 TTCTCTGGTGAAGCAAAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA 384  
  
Qy 241 CTTGCTCTCCGGGACAGCAGCTGGGAGATGAGA 274  
Db 385 CTTGCTCTCCGGGACAGCAGCTGGGAGATGAGA 418

RESULT 9  
AAH64901  
ID AAH64901 standard; cdna; 1361 BP.  
AC AAH64901;  
XX  
XX  
DT 11-SEP-2001 (first entry)  
XX  
XX Human secreted protein cdna, SEQ ID NO: 177.  
DE  
DE Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;  
KW GENSET; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200142451-A2.  
PN  
XX  
XX 14-JUN-2001.  
PD  
XX  
XX 07-DEC-2000; 2000WO-IB01938.  
PF  
XX  
XX 08-DEC-1999; 9905-0169629.  
PR  
XX  
XX 06-MAR-2000; 2000US-0187470.  
PR  
XX  
XX (GEST ) GENSET.  
PA  
XX  
XX Dumas Milne Edwards J, Bouqueleret L, Jobert S;  
PI  
XX  
XX WPI; 2001-367870/38.  
DR  
XX  
XX P-PSDB; AAG89298.  
DR  
XX  
XX Full length GENSET human nucleic acids encoding potentially secreted  
PT proteins, useful in gene therapy and vaccination against a variety of  
PT diseases, and for diagnosis of those diseases -  
XX  
XX Claim 7; Page 741-742; 921pp; English.  
PS  
XX  
XX The invention relates to full length GENSET human nucleic acids encoding  
CC potentially secreted proteins. The nucleic acids and the polypeptides

CC they encode may be used in the prevention, treatment and diagnosis of  
CC diseases associated with inappropriate GENSET gene expression. For  
CC example, they be used to treat disorders associated with decreased  
CC GENSET gene expression by rectifying mutations or deletions in a  
CC patient's genome that affect the activity of GENSET or by supplementing  
CC the patient's own production of GENSET polypeptides. Conversely,  
CC antisense nucleic acid molecules may be administered to down regulate  
CC GENSET expression by binding with the cells' own genes and preventing  
CC their expression. The sense and antisense nucleic acids may also be  
CC used as DNA probes in diagnostic assays to detect and quantitate the  
CC presence of similar nucleic acid sequences in samples, and hence to  
CC determine which patients may be in need of restorative therapy.  
CC The GENSET polypeptides may be used as antigens in the production of  
CC antibodies and in assays to identify modulators (agonists and  
CC antagonists) of GENSET polypeptide expression and activity. The  
CC present sequence is a GENSET nucleic acid of the invention.  
XX

SO Sequence 1361 BP; 438 A; 245 C; 302 G; 375 T; 1 other;  
  
Query Match 6.1%; Score 149; DB 22; Length 1361;  
Best Local Similarity 100.0%; Pred. No. 5.3e-62;  
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 2299 TTCTAAATCCTGCCCTCTGTTCCTTTCTCTCTTCTGAAAGTCCACACACCATTTCTGTCC 2358  
Db 1177 TTCTAAATCCTGCCCTCTGTTCCTTTCTCTCTTCTGAAAGTCCACACACCATTTCTGTCC 1236  
  
Qy 2359 TTCCCCAGTTTCTCGCCCTCCACCCCTCCAGCTTCATGCTGCTGCTTAATAAAA 2418  
Db 1237 TTCCCCAGTTTCTCGCCCTCCACCCCTCCAGCTTCATGCTGCTGCTTAATAAAA 1296  
  
Qy 2419 ATGGACATATTTTCTCTATAAAAAAAA 2447  
Db 1297 ATGGACATATTTTCTCTATAAAAAAAA 1325

RESULT 10  
ABL99884  
ID ABL99884 standard; cdna; 373 BP.  
AC ABL99884;  
XX  
XX  
DT 03-OCT-2002 (first entry)  
XX  
XX Human secretory polynucleotide (sptm) 139.  
DE  
DE  
XX  
XX Human; ss; gene; secretory protein; secretory polynucleotides; SPTM;  
KW SPTM-related disease; somatic gene therapy; germline gene therapy;  
KW severe combined immunodeficiency; intracellular parasite protection;  
KW fungal parasite; protozoan parasite; cell proliferative disorder; cancer;  
KW immune disorder; AIDS; neurological disorder; Parkinson's disease;  
KW motor neuron disorder; demyelinating disease; multiple sclerosis;  
KW meningitis; abscess; prion diseases; cerebral palsy;  
KW neuroskeletal disorder; peripheral nervous system disorder;  
KW dermatomyositis; polymyositis; myopathy; myasthenia gravis;  
KW mental disorder; Tourette's syndrome.  
XX  
OS  
XX Homo sapiens.  
PN  
XX  
XX WO200220756-A2.  
PD  
XX  
XX 14-MAR-2002.  
XX  
XX 30-AUG-2001; 2001WO-US27297.  
PF  
XX  
XX 05-SEP-2000; 2000US-229747P.  
PR  
XX  
XX 05-SEP-2000; 2000US-229748P.  
PR  
XX  
XX 05-SEP-2000; 2000US-229749P.  
PR  
XX  
XX 05-SEP-2000; 2000US-229750P.  
PR  
XX  
XX 05-SEP-2000; 2000US-229751P.  
PR  
XX  
XX 05-SEP-2000; 2000US-230016P.  
PR  
XX  
XX 06-SEP-2000; 2000US-230583P.  
PR  
XX  
XX 06-SEP-2000; 2000US-230505P.

```
PR 06-SEP-2000; 2000US-230514P.
PR 06-SEP-2000; 2000US-230515P.
PR 06-SEP-2000; 2000US-230517P.
PR 06-SEP-2000; 2000US-230518P.
PR 06-SEP-2000; 2000US-230519P.
PR 06-SEP-2000; 2000US-230595P.
PR 06-SEP-2000; 2000US-230596P.
PR 06-SEP-2000; 2000US-230597P.
PR 06-SEP-2000; 2000US-230599P.
PR 06-SEP-2000; 2000US-230610P.
PR 06-SEP-2000; 2000US-230864P.
PR 06-SEP-2000; 2000US-230865P.
PR 06-SEP-2000; 2000US-230888P.
PR 06-SEP-2000; 2000US-230899P.
PR 06-SEP-2000; 2000US-230990P.
PR 06-SEP-2000; 2000US-230896P.
PR 07-SEP-2000; 2000US-230897P.
PR 07-SEP-2000; 2000US-230951P.
PR 07-SEP-2000; 2000US-231163P.
PR 07-SEP-2000; 2000US-231832P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
XX Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;
PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
PI Momiya MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
PI Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Dafo A;
PI Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;
XX
XX WPI: 2002-315658/35.
XX P-PSDB; ABB97887.
XX
XX Polynucleotide sequences encoding human secretory proteins useful for
PT gene therapy of e.g. genetic deficiency disorders, cancers, and
PT diseases caused by intracellular parasites -
XX
XX Claim 1; Page 331; 585pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
CC secretory (SPM) proteins. The SPM DNA and amino acid sequences are
CC useful for treating a disease or condition associated with the expression
CC of functional SPM. The SPM DNA sequences are useful for somatic or
CC germline gene therapy to correct a genetic deficiency (e.g. severe
CC combined immunodeficiency). The SPM DNA sequences are also useful in
CC providing protection against intracellular parasites (e.g. fungal
CC parasites and protozoan parasites). The SPM DNA and protein sequences
CC are also useful for diagnosing cell proliferative disorders, cancer,
CC immune disorders (e.g. AIDS), neurological disorders (e.g. Parkinson's
CC disease), motor neuron disorders, demyelinating diseases (e.g. multiple
CC sclerosis), meningitis, abscesses, prion diseases, cerebral palsy,
CC neuroskeletal disorders, peripheral nervous system disorders,
CC dermatomyositis and polymyositis, myopathy, myasthenia gravis, and mental
CC disorders (e.g. Tourette's syndrome). cDNA sequences ABL99746 - ABL99929
CC represent human secretory polynucleotides of the invention.
XX
XX Sequence 373 BP; 92 A; 84 C; 98 G; 99 T; 0 other;
SQ
Query Match 4.4%; Score 107; DB 24; Length 373;
Best Local Similarity 99.4%; Pred. No. 1.5e-41;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 39 GGCATCTCTGAGAAGCTTCAACTCTGGAGGCAATGGGTGGAAGGAAGATGACTGCAG 98
DB 205 GGCATCTCTGAGAAGCTTCAACTCTGGAGGCAATGGGTGGAAGGAAGATGACTGCAG 264
QY 99 TTCTGTGAAGAACAGACACACATCCGGAAACCTTCATTTTATGGAAGTCTGGG 158
DB 265 TTCTGTGTGAAGAACAGACACACATCCGGAAACCTTCATTTTATGGAAGTCTGGG 324
QY 159 ATCAGGAGCTTTCCTCAGAAGTTTTTCCTGGTGAAGCAA 196
DB 325 ATCAGGAGCTTTCCTCAGAAGTTTTTCCTGGTGAAGCAA 362
```

## RESULT 11

```
ABN35390
ID ABN35390 standard; DNA; 60 BP.
XX
XX AC ABN35390;
XX
XX DT 15-JUL-2002 (first entry)
XX
XX DE Human spliced transcript detection oligonucleotide SEQ ID NO:8138.
XX
XX KW Human; mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.
XX
XX OS Homo sapiens.
XX
XX PN WC200210449-A2.
XX
XX PD 07-FEB-2002.
XX
XX PF 20-JUL-2001; 2001WO-1B01903.
XX
XX PR 28-JUL-2000; 2000US-221607P.
XX PR 02-MAY-2001; 2001US-287724P.
XX
XX PA (COMP-) COMPUGEN INC.
XX
XX PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
XX WPI: 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
XX Example 1; SEQ ID 8138; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 60 BP; 13 A; 18 C; 10 G; 19 T; 0 other;
SQ
Query Match 2.5%; Score 60; DB 24; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.1e-18;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1526 TGCAATTTTCAGGAGACATATTCACCTCCCTGCTCTTCCAAACCTGGTGTCTATCGGC 1585
DB 1 TGCAATTTTCAGGAGACATATTCACCTCCCTGCTCTTCCAAACCTGGTGTCTATCGGC 60
```

```
RESULT 12
AAS70157
ID AAS70157 standard; cDNA; 1512 BP.
XX
AC AAS70157;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #5961.
XX
Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
PI WPI; 2001-639362/73.
XX
DR P-PSDB; ABG05970.
XX
New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 5961; 103pp; English.
XX
The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1512 BP; 343 A; 391 C; 470 G; 308 T; 0 other;

Query Match 1.3%; Score 32; DB 23; Length 1512;
Best Local Similarity 100.0%; Pred. No. 5e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 CCTACAGCAAGGCTGTGGATTGCTGGTCCAT 683
|
DB 598 CCTACAGCAAGGCTGTGGATTGCTGGTCCAT 629

RESULT 13
```

```
AAS68767
ID AAS68767 standard; cDNA; 1774 BP.
XX
AC AAS68767;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #4571.
XX
Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
PI WPI; 2001-639362/73.
XX
DR P-PSDB; ABG04580.
XX
New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 4571; 103pp; English.
XX
The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1774 BP; 421 A; 488 C; 477 G; 388 T; 0 other;

Query Match 1.3%; Score 32; DB 23; Length 1774;
Best Local Similarity 100.0%; Pred. No. 5e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 CCTACAGCAAGGCTGTGGATTGCTGGTCCAT 683
|
DB 995 CCTACAGCAAGGCTGTGGATTGCTGGTCCAT 1026

RESULT 14
AAS90919
ID AAS90919 standard; cDNA; 1774 BP.
```

```
XX AC AAS90919;
XX DT 13-FEB-2002 (first entry)
XX DE
XX KW DNA encoding novel human diagnostic protein #26723.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
XX P-PSDB; ABG26732.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PS Claim 1; SEQ ID No 26723; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human
XX CC diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 1774 BP; 421 A; 488 C; 477 G; 388 T; 0 other;

Query Match 1.3%; Score 32; DB 23; Length 1774;
Best Local Similarity 100.0%; Pred. No. 5e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 CCCTACAGCAAGCGTGTGGATTGCTGTCAT 683
Db 995 CCCTACAGCAAGCGTGTGGATTGCTGTCAT 1026

RESULT 15
AAD36148/c
ID AAD36148 standard; DNA; 25 BP.
XX AC AAD36148;
```

```
XX DT 09-AUG-2002 (first entry)
XX DE Human CAMK-XI full length cDNA constructing 5' RACE PCR primer.
XX KW Human; cytostatic; antisense gene therapy; screening; protein kinase;
XX KW cancer; liver; colon; tumour; inflammation; arthritic synovium; CAMK-XI;
XX KW calmodulin kinase; enzyme; RACE; rapid amplification of cDNA end; primer;
XX KW PCR; ss.
XX OS Homo sapiens.
XX PN WO200224947-A2.
XX PD 28-MAR-2002.
XX PF 20-SEP-2001; 2001WO-IB02237.
XX PR 20-SEP-2000; 2000US-233999p.
XX PR 02-OCT-2000; 2000US-237419p.
XX PR 02-OCT-2000; 2000US-237423p.
XX PR 04-OCT-2000; 2000US-238558p.
XX PR 10-MAY-2001; 2001US-290555p.
XX PA (KINE-) KINETEK PHARM INC.
XX PA (UYER-) UNIV BRITISH COLUMBIA.
XX PI Yoganathan T, Delaney AD;
XX WPI: 2002-394145/42.
XX PT Diagnosing cancer, comprises determining the upregulation of expression
XX PT of a nucleic acid sequence encoding a protein kinase or upregulation of
XX PT expression of the protein kinase, in the cancer -
XX PS Example 2; Page 82; 87pp; English.
XX CC The invention relates to a method for screening biologically active agent
XX CC that modulates cancer associated protein kinase function. The invention
XX CC also relates to a method for diagnosing cancer comprising determining the
XX CC upregulation of expression of a nucleic acid sequence encoding a protein
XX CC kinase. The method is useful for diagnosing cancer. A protein kinase is
XX CC useful for screening biological agents that modulate cancer associated
XX CC protein kinase function. Downregulating the activity of protein kinase is
XX CC useful for inhibiting the growth of a cancer cell, e.g. liver or colon
XX CC cancer. A nucleic acid encoding protein kinase is useful to screen biopsy
XX CC derived tumours and inflammatory samples such as arthritic synovium, for
XX CC amplified DNA in the cell or increased expression of corresponding mRNA
XX CC or protein and is also useful to detect differences in expression levels
XX CC such as molecular weight, amino acid and nucleotide sequences between the
XX CC two cells. The present sequence is 5' rapid amplification of cDNA ends
XX CC (RACE) PCR primer used for constructing human calmodulin kinase CAMK-XI
XX CC full length cDNA.
XX CC Note: The present sequence referred as SEQ ID NO: 17 is also shown in
XX CC page 25 of the specification but lacks two bases at the 5' end.
XX SQ Sequence 25 BP; 7 A; 2 C; 14 G; 2 T; 0 other;

Query Match 1.0%; Score 25; DB 24; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2358 CTTCCCCAGTTTCTCGCCCTCCAC 2382
Db 25 CTTCCCCAGTTTCTCGCCCTCCAC 1

Search completed: March 14, 2003, 17:35:56
Job time : 573 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: March 14, 2003, 15:22:10 ; Search time 7161 Seconds  
(without alignments)  
9944.779 Million cell updates/sec  
Title: US-09-960-643-1  
Perfect score: 2447  
Sequence: 1 tggagtgggagctcaagcag.....tttttcttaaaaaaaaaa 2447  
Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0  
Searched: 2054640 seqs, 14551402878 residues  
Word size : 0  
Total number of hits satisfying chosen parameters: 4109280  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2447	100.0	2447	6	AX399682 Sequence
2	2396	97.9	2474	9	BC032787 Homo sapi
3	2143	87.6	2612	9	AK095713 Homo sapi
4	2048	83.7	2464	9	AF428261 Homo sapi
5	1636	66.9	1738	9	HS272L161
6	903	36.9	157875	9	HS272L16 Human gen
7	78	3.2	1249	9	HUMDIS70R
8	71	2.9	2416	10	BC021840 Mus muscu
9	71	2.9	2427	10	AF428262 Mus muscu
10	62	2.5	182054	10	AL365314 Mouse DNA
11	62	2.5	184754	2	AC022675 Mus muscu
12	59	2.4	1013	10	D86557
13	40	1.6	178272	2	AC126166 Rattus norv
14	40	1.6	191584	2	AC120477 Rattus no
15	32	1.3	1448	10	BC014825 Mus muscu
16	32	1.3	1480	9	HUMCKI
17	32	1.3	2689	9	AK094026 Homo sapi
18	32	1.3	155313	2	AC068315 Homo sapi
19	32	1.3	157007	2	AC066599 Homo sapi
20	32	1.3	161903	2	AC021996 Homo sapi
21	32	1.3	185061	9	AC022382 Homo sapi
22	32	1.3	189430	2	AC011610 Homo sapi
23	32	1.3	197360	2	AC018829 Homo sapi
24	27	1.1	191584	2	AC120477 Rattus no
25	25	1.0	25	6	AX399696 Sequence
26	25	1.0	25	6	AX399703 Sequence
27	25	1.0	30	6	AX399708 Sequence
28	25	1.0	30	6	AX399709 Sequence
29	25	1.0	31	6	AX399704 Sequence
30	24	1.0	30	6	AX399710 Sequence
31	23	0.9	296	6	AR050391 Sequence
32	23	0.9	534	6	AX024570 Sequence
33	23	0.9	3380	6	AR078528 Sequence
34	23	0.9	3380	9	HSFAN
35	23	0.9	146872	2	AC128229 Rattus no
36	23	0.9	167898	2	AC103528 Rattus no
37	22	0.9	755	6	AR062879 Sequence
38	22	0.9	824	9	HSGRSPROT
39	22	0.9	829	9	BC016281 Homo sapi
40	22	0.9	860	9	HS325898 Homo sapi
41	22	0.9	1420	9	AK001010 Homo sapi
42	22	0.9	2160	10	MM021392 Mus musculu
43	22	0.9	2347	9	AK055324 Homo sapi
44	22	0.9	2462	9	AK097626 Homo sapi
45	22	0.9	2648	10	BC031180 Mus muscu

ALIGNMENTS

RESULT 1  
AX399682  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

AX399682  
Sequence 3 from Patent WO0224947.  
AX399682  
AX399682.1 GI:21335455  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Delaney,A.D. and Yoganathan,T.  
Cancer associated protein kinases and their uses  
Patent: WO 0224947-A 3 28-MAR-2002;

AX399682  
Sequence 3 from Patent WO0224947.  
AX399682  
AX399682.1 GI:21335455  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Delaney,A.D. and Yoganathan,T.  
Cancer associated protein kinases and their uses  
Patent: WO 0224947-A 3 28-MAR-2002;

AX399682  
Sequence 3 from Patent WO0224947.  
AX399682  
AX399682.1 GI:21335455  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Delaney,A.D. and Yoganathan,T.  
Cancer associated protein kinases and their uses  
Patent: WO 0224947-A 3 28-MAR-2002;

FEATURES	KINETEK PHARMACEUTICALS INC (CA); UNIV BRITISH COLUMBIA (CA)									
	source	Location/Qualifiers 1..2447 /organism="Homo sapiens" /db_xref="taxon:9606" 70..1500 /note="unnamed protein product" /codon_start=1 /protein_id="CAD33521.1" /db_xref="GI:21335456"								
CDS	TRANSLATION="MKRKEEDDCSSWKQTNIRKTFIPMEVLGGSPFSEVLVKQRL TGLFKLCIKKSPAFRDSLENIKIKHENIVITLEDIYESTHYLYVMQLVSG GELFDRLGRVYTEKDSLIVQOVLKSAVKYLHENGIVHRDLKPNLLYLTPEENSKI MITDFGLSKWEONGINSTAGCTPGYVAPVLAQPKYSKAVDCWSIGVITYILICGVPP PYETRSKLFPEKIEGYYPESFPDDISESAKDFICHLEKDPNERVTCERALSHPW IDGNTALHREDIYFVSQIQKNFAKSKWRQAFNAAVVHMKRLKLNHLSPGVREVE NRPPEQASTSPSPSEITTEAPVLDSHVALPALTOIPQCHRRPTAPGGRSLNCL VNGSLHISSSLVPMHOGSLAAGPCGCCSCLNITGSKGSSYCSSEPTLLKANKKNFK SEVMVPVKASGSHCRAGOTGVCLIM"									
BASE COUNT	590 a	707 c	604 g	546 t						
ORIGIN										
	Query Match	100.0%;	Score 2447;	DB 6;	Length 2447;					
	Best Local Similarity	100.0%;	Pred. No. 0;							
	Matches 2447;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;					
Qy	1	TGGAGTGGGAGCTCAAGCAGGATTCTTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT	60							
Db	1	TGGAGTGGGAGCTCAAGCAGGATTCTTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT	60							
Qy	61	CTGGAGGCAATGGGTTCGAAAGAAAGATGACTGCAAGTTCCTCGAAGAAACAGACCAACC	120							
Db	61	CTGGAGGCAATGGGTTCGAAAGAAAGATGACTGCAAGTTCCTCGAAGAAACAGACCAACC	120							
Qy	121	AACATCCGGAAAAACCTTCATTTTATGGAAGTGTGGAGTGTGGAGCTTCTCAGAAGTT	180							
Db	121	AACATCCGGAAAAACCTTCATTTTATGGAAGTGTGGAGTGTGGAGCTTCTCAGAAGTT	180							
Qy	181	TTCTCTGGTGAAGCAAGACTGACATGGAAGCTCTTTGCTCTGAAAGTGCATCAAGAAGTCA	240							
Db	181	TTCTCTGGTGAAGCAAGACTGACATGGAAGCTCTTTGCTCTGAAAGTGCATCAAGAAGTCA	240							
Qy	241	CCTGCCTTCGCGGACAGAGCCTGGAGAATGAGATTGCTGTGTTGAAAAAGATCAAGCAT	300							
Db	241	CCTGCCTTCGCGGACAGAGCCTGGAGAATGAGATTGCTGTGTTGAAAAAGATCAAGCAT	300							
Qy	301	GAACATATTGTGACCTGGAGACATCTATGAGAGCACCCACTACTACTCTGGTCTATG	360							
Db	301	GAACATATTGTGACCTGGAGACATCTATGAGAGCACCCACTACTACTCTGGTCTATG	360							
Qy	361	CAGCTTGTCTTGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAG	420							
Db	361	CAGCTTGTCTTGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAG	420							
Qy	421	AAGGATGCCAGTCTGGTATCCAGAGCTTTGTTCGGAGTGAATACCTACATCAGAAT	480							
Db	421	AAGGATGCCAGTCTGGTATCCAGAGCTTTGTTCGGAGTGAATACCTACATCAGAAT	480							
Qy	481	GGCATCGCCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC	540							
Db	481	GGCATCGCCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC	540							
Qy	541	TCTAAGATCATGATCATGACTTTGGTCTGTCCAAGATGGAAACAGAAATGGCATCATGCC	600							
Db	541	TCTAAGATCATGATCATGACTTTGGTCTGTCCAAGATGGAAACAGAAATGGCATCATGCC	600							
Qy	601	ACTGCCTGTGGACCCAGGCTAGTGGCTCCAGAAGTGTGGCCCGCAGAAACCTTACAGC	660							
Db	601	ACTGCCTGTGGACCCAGGCTAGTGGCTCCAGAAGTGTGGCCCGCAGAAACCTTACAGC	660							
Qy	661	AAGGCTGTGGATTGCTGGTCCATCGGCGTCAATCACTACATATTGCTGTGGATACCCC	720							
Db	661	AAGGCTGTGGATTGCTGGTCCATCGGCGTCAATCACTACATATTGCTGTGGATACCCC	720							

Qy	721	CCGTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTTACTATGAG	780	
Db	721	CCGTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTTACTATGAG	780	
Qy	781	TTTGAGTCTCCATTCTGGGATGACATTTCTGAGTCAGCGCAAGGACTTTATTTTGCCACTTG	840	
Db	781	TTTGAGTCTCCATTCTGGGATGACATTTCTGAGTCAGCGCAAGGACTTTATTTTGCCACTTG	840	
Qy	841	CTTGAGAAGGATTCGGAACAGCGGTACACCTGTGAGAAGGCCTTGAGTCATCCCTGGATT	900	
Db	841	CTTGAGAAGGATTCGGAACAGCGGTACACCTGTGAGAAGGCCTTGAGTCATCCCTGGATT	900	
Qy	901	GACGGAACACCGCCCTCCACCGGACATCTACCCATCAGTCAGCCTCCAGATCCAGAG	960	
Db	901	GACGGAACACCGCCCTCCACCGGACATCTACCCATCAGTCAGCCTCCAGATCCAGAG	960	
Qy	961	AACTTTGCTAAGAGCAAGTGGAGGCAAGCTTCAAGCGCAGCAGCTGTGGTGACCAATG	1020	
Db	961	AACTTTGCTAAGAGCAAGTGGAGGCAAGCTTCAAGCGCAGCAGCTGTGGTGACCAATG	1020	
Qy	1021	AGGAAGCTACACATGAACCTGCACAGCCCGCGCTCCGCCAGAGGTGAGAACAGGCCG	1080	
Db	1021	AGGAAGCTACACATGAACCTGCACAGCCCGCGCTCCGCCAGAGGTGAGAACAGGCCG	1080	
Qy	1081	CCTGAAACTCAAGCCTTCAGAAACCTCTAGACCCAGCTCCCTGAGATCAACCATCCAG	1140	
Db	1081	CCTGAAACTCAAGCCTTCAGAAACCTCTAGACCCAGCTCCCTGAGATCAACCATCCAG	1140	
Qy	1141	GCACCTGTCTGGACACAGTGTAGCAGTCCCTGCCCTGACCCAAATACCTTGCAGCAT	1200	
Db	1141	GCACCTGTCTGGACACAGTGTAGCAGTCCCTGCCCTGACCCAAATACCTTGCAGCAT	1200	
Qy	1201	GGCGCGCGCCACTGCGCTGTGGCAGTCCCTCAACTGCTGTGTCATTAATGGCTCCCTC	1260	
Db	1201	GGCGCGCGCCACTGCGCTGTGGCAGTCCCTCAACTGCTGTGTCATTAATGGCTCCCTC	1260	
Qy	1261	CACATCAGCAGCAGCCTGTGCTGATCATGAGGGTCCCTGGCGCGCGCGCTGTGCG	1320	
Db	1261	CACATCAGCAGCAGCCTGTGCTGATCATGAGGGTCCCTGGCGCGCGCGCTGTGCG	1320	
Qy	1321	TGCTGTCTCAGCTGCTGAACATTTGGAGCAAGGAAAGTCTCTTACTGCTTGAGCCC	1380	
Db	1321	TGCTGTCTCAGCTGCTGAACATTTGGAGCAAGGAAAGTCTCTTACTGCTTGAGCCC	1380	
Qy	1381	ACACTCCTCAAAAAGCCCAACAAAAACAGACTTCAAGTCGAGGATCATGTTACAGTT	1440	
Db	1381	ACACTCCTCAAAAAGCCCAACAAAAACAGACTTCAAGTCGAGGATCATGTTACAGTT	1440	
Qy	1441	AAAGCCAGTGGCAGCTCCACCTGCCCGGAGGCGAGCTGGAGTCTGTCTCATTTATGTA	1500	
Db	1441	AAAGCCAGTGGCAGCTCCACCTGCCCGGAGGCGAGCTGGAGTCTGTCTCATTTATGTA	1500	
Qy	1501	TTCTCTGAGCCTGTGCTATGTCACTGCAATTTTTCAGGAGACATATTCAACTCTCTGCT	1560	
Db	1501	TTCTCTGAGCCTGTGCTATGTCACTGCAATTTTTCAGGAGACATATTCAACTCTCTGCT	1560	
Qy	1561	CTTCCAAAACCTGTGCTATCTCCGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1620	
Db	1561	CTTCCAAAACCTGTGCTATCTCCGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1620	
Qy	1621	AGCAGAGCAGTTCCTGGCCAGAGCAGCAGCTGCTGCCAGGGGCGAGCCCTCATAG	1680	
Db	1621	AGCAGAGCAGTTCCTGGCCAGAGCAGCAGCTGCTGCCAGGGGCGAGCCCTCATAG	1680	
Qy	1681	GAGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1740	
Db	1681	GAGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1740	
Qy	1741	GGTGGCCACAGCTTCCAGGCTTCCCTGACCTGCTGCTATGCCCCACACCCCTACGTTG	1800	
Db	1741	GGTGGCCACAGCTTCCAGGCTTCCCTGACCTGCTGCTATGCCCCACACCCCTACGTTG	1800	

```
Qy 1801 CCGTGGCTCTGTCGAGGTGACGTACGTCTCGCCTGGGTCTGTGCTGTGCTGTA 1860
|||||
Db 1801 CCGTGGCTCTGTCGAGGTGACGTACGTCTCGCCTGGGTCTGTGCTGTGCTGTA 1860
|||||
Qy 1861 AAAGCTTAATGGGCTGCCAGGTGTGTACCTTCTCCAGCAAGCAACCATATGGAGCANC 1920
|||||
Db 1861 AAAGCTTAATGGGCTGCCAGGTGTGTACCTTCTCCAGCAAGCAACCATATGGAGCANC 1920
|||||
Qy 1921 TACCCAGACGCCACTCTGCACACACTCACTCCACCTCTCAAGCCTCCAACTCTTGCC 1980
|||||
Db 1921 TACCCAGACGCCACTCTGCACACACTCACTCCACCTCTCAAGCCTCCAACTCTTGCC 1980
|||||
Qy 1981 CAGATTGGGCTCATTAATGTCGTTGCTGCGCATCTGCAATGACAGCGACTCCCA 2040
|||||
Db 1981 CAGATTGGGCTCATTAATGTCGTTGCTGCGCATCTGCAATGACAGCGACTCCCA 2040
|||||
Qy 2041 TGGTGGTCTGCTGAGCTCTTCAAGTTCCTTAATCTTAACCTCAGGATTAGTCCCAAG 2100
|||||
Db 2041 TGGTGGTCTGCTGAGCTCTTCAAGTTCCTTAATCTTAACCTCAGGATTAGTCCCAAG 2100
|||||
Qy 2101 TGCCTGAGACCCAGCAGCAGACTTCTGCGCTCTTCCCTGCTCAATCTAAAGCAGT 2160
|||||
Db 2101 TGCCTGAGACCCAGCAGCAGACTTCTGCGCTCTTCCCTGCTCAATCTAAAGCAGT 2160
|||||
Qy 2161 GCCACCCCTCCAAAGTGAATGAAGAAGTTCATGAGTAAAGGCTGCAAGGAATCTT 2220
|||||
Db 2161 GCCACCCCTCCAAAGTGAATGAAGAAGTTCATGAGTAAAGGCTGCAAGGAATCTT 2220
|||||
Qy 2221 ATCTGCGCCACATCTCTCGTGCACACACCCCAATGAGTAACTTGAAGTTGACTAT 2280
|||||
Db 2221 ATCTGCGCCACATCTCTCGTGCACACACCCCAATGAGTAACTTGAAGTTGACTAT 2280
|||||
Qy 2281 TTTAATGTCGCCAGGAGTCTAATCTGCTCTGTTCCCTCTGTTCTCTCTGTAAGTCC 2340
|||||
Db 2281 TTTAATGTCGCCAGGAGTCTAATCTGCTCTGTTCCCTCTGTTCTCTCTGTAAGTCC 2340
|||||
Qy 2341 AGCACACCATCTTGTCTTCCCTCCAGTTCTCTCGCCTCCACCCCTCCAGCTTCATGCTC 2400
|||||
Db 2341 AGCACACCATCTTGTCTTCCCTCCAGTTCTCTCGCCTCCACCCCTCCAGCTTCATGCTC 2400
|||||
Qy 2401 AGTGTGTGCTTAATAAATGGACATATTTTCTCTTAAAAAAA 2447
|||||
Db 2401 AGTGTGTGCTTAATAAATGGACATATTTTCTCTTAAAAAAA 2447
|||||

RESULT 2
BC032787
LOCUS BC032787 2474 bp mRNA linear PRI 27-JUN-2002
DEFINITION Homo sapiens, calcium/calmodulin-dependent protein kinase IG, clone
MGC:44894 IMAGE:5179957, mRNA, complete cds.
ACCESSION BC032787
VERSION BC032787.1 GI:21619664
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2474)
Strausberg,R.
Direct Submission
Submitted (06-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
```

Web site: <http://www.nisc.nih.gov/>  
Contact: nisc\_mgc@nih.gov  
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
Maduro, Q.L., Masliello, C., Maskeri, B., Mastriano, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
Tsurgoun, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 68 Row: K Column: 17  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 1419644.

FEATURES  
Location/Qualifiers  
1..2474  
/organism="Homo sapiens"  
/db\_xref="LocusID:57172"  
/clone="MGC:44894 IMAGE:5179957"  
/tissue\_type="Brain, Lung, Testis, adult, pooled whole"  
/clone\_lib="NIH\_MGC\_115"  
/lab\_host="DH105"  
/note="Vector: pCMV-SPORT6"  
70..1500  
/codon\_start=1  
/product="calcium/calmodulin-dependent protein kinase IG"  
/protein\_id="AAH32787.1"  
/db\_xref="GI:21619665"  
/translation="MGRKREDCCSWKKOTNIRKTFIFMVLGSCAFSEFLVKORL  
TGLFLALCKIKSPAPRDSLENEIAVLAKIKHENIVLTEDIYESTHYLVQVLVSG  
GELFDRLIRGVYETKDALVLIQVLSAVKYDHENGIVHRDLKPNLLYLTPTPEENSKI  
MITDFGLSKMEQNGINSTAGTGYVAPVLAQKPYSCAVDCNMGITVYIILLGYP  
FYDETATSKLFEIKEGYEFSEPFNDI2SEKADFICHLLEKDPNRYTCERKLSHPW  
IDGNTALHRDIYPSVSIQOKNFARNAAVVHMKRKLHMLHLSHPCGVREVE  
NRPEPTQASSETSPSPSEITITAPVLDHVALPALTPOLPCOHGRRPTAPGGRSLNCL  
VNGSLHSSSLVPMHOGSLAAGCGCCSCLNIGSKGKSSYSEPTLLKKANKQNFK  
SEVMFPVKAGSSSHCRAGTGVCCLIM"

BASE COUNT 618 a 707 c 603 g 546 t  
ORIGIN  
Query Match 97.9%; Score 2396; DB 9; Length 2474;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2446; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TGGAGTGGGAGCTCAAGCAGGATTCTTCCCGAGTCCCTGGCATCTCTCAGAGCTTCAACT 60  
|||||  
Db 1 TGGAGTGGGAGCTCAAGCAGGATTCTTCCCGAGTCCCTGGCATCTCTCAGAGCTTCAACT 60  
|||||  
Qy 61 CTGGAGGCAATGGTCCGAAGGAAGATGACTGCAGTTCCTTGGGAAGAACAGACACC 120  
|||||  
Db 61 CTGGAGGCAATGGTCCGAAGGAAGATGACTGCAGTTCCTTGGGAAGAACAGACACC 120  
|||||  
Qy 121 AACATCCGGAAAACTTCATTTTATGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTT 180  
|||||  
Db 121 AACATCCGGAAAACTTCATTTTATGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTT 180  
|||||  
Qy 181 TTCCTGTTGAAGCAAGACTGACTGGGAAGCTCTTCTCTGAAGTGCATCAAGAAGTCA 240  
|||||  
Db 181 TTCCTGTTGAAGCAAGACTGACTGGGAAGCTCTTCTCTGAAGTGCATCAAGAAGTCA 240  
|||||  
Qy 241 CTGCTCTTCCGGGACAGCAGCTCTGGGAATGAGATTGCTGTGTGTAAGAAAGATCAAGCAT 300  
|||||  
Db 241 CTGCTCTTCCGGGACAGCAGCTCTGGGAATGAGATTGCTGTGTGTAAGAAAGATCAAGCAT 300  
|||||  
Qy 301 GAAACATTTGACCCCTGGAGGACATCTATGAGAGCACACCCTACTACTACTACCTGGTCTG 360  
|||||  
Db 301 GAAACATTTGACCCCTGGAGGACATCTATGAGAGCACACCCTACTACTACTACCTGGTCTG 360  
|||||  
Qy 361 CAGCTTGTCTGTTGGGAGCTCTTTACCGGATCTTGGAGCGGCTGCTACACAGAG 420  
|||||

Db 361 CAGCTTTCTTGGTGGGAGCTCTTTGACCGGATCCTGGAGGGGGTGTCTACACAGAG 420  
Qy 421 AAGGATGCCAGTCTGFGATCCACAGAGTCTTGTGCGCAGTGAATACCTACATGAGAT 480  
Db 421 AAGGATGCCAGTCTGFGATCCACAGAGTCTTGTGCGCAGTGAATACCTACATGAGAT 480  
Qy 481 GGCAATCTCCACAGAGACTTAAAGCCGGAACCTGCTTACCTTACCCCTGAAGAGAC 540  
Db 481 GGCAATCTCCACAGAGACTTAAAGCCGGAACCTGCTTACCTTACCCCTGAAGAGAC 540  
Qy 541 TCTAAGATCATGATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGSCATCATGTCC 600  
Db 541 TCTAAGATCATGATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGSCATCATGTCC 600  
Qy 601 ACTGCTGTGGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCCGAGAAACCCCTACAGC 660  
Db 601 ACTGCTGTGGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCCGAGAAACCCCTACAGC 660  
Qy 661 AAGGCTGTGGATTGCTGGTCCATCGGCTCATCACCTACATATTGCTGTGGATACCC 720  
Db 661 AAGGCTGTGGATTGCTGGTCCATCGGCTCATCACCTACATATTGCTGTGGATACCC 720  
Qy 721 CCCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTACTATGAG 780  
Db 721 CCCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTACTATGAG 780  
Qy 781 TTTGAGTCTCCATCTGGGATGACATTTCTGAGTCAAGCAAGGACTTATTGCGCACTTG 840  
Db 781 TTTGAGTCTCCATCTGGGATGACATTTCTGAGTCAAGCAAGGACTTATTGCGCACTTG 840  
Qy 841 CTTGAGAAGGATCCGACAGCGGTACACCTGTGAGAAGGCCCTTGAGTCATCCCTGGATT 900  
Db 841 CTTGAGAAGGATCCGACAGCGGTACACCTGTGAGAAGGCCCTTGAGTCATCCCTGGATT 900  
Qy 901 GACGGAACAGCGCTCCACCGGACATCTACCCATCAGTCAGCTCAGCTCCAGATCCAGAG 960  
Db 901 GACGGAACAGCGCTCCACCGGACATCTACCCATCAGTCAGCTCAGCTCCAGATCCAGAG 960  
Qy 961 AACTTTGCTAAGACAAAGTGGAGCAAGCCTTCAAGCGAGCAGCTGTGGTGCACCAATG 1020  
Db 961 AACTTTGCTAAGACAAAGTGGAGCAAGCCTTCAAGCGAGCAGCTGTGGTGCACCAATG 1020  
Qy 1021 AGGAGCTACACATGACCTGACAGCGCGGGCTCGGCCAGAGGTGAGAACAGCGCG 1080  
Db 1021 AGGAGCTACACATGACCTGACAGCGCGGGCTCGGCCAGAGGTGAGAACAGCGCG 1080  
Qy 1081 CCTGAAACTCAAGCCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCACCATCCAGAG 1140  
Db 1081 CCTGAAACTCAAGCCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCACCATCCAGAG 1140  
Qy 1141 GCACCTGTCTGGACACAGTGTAGCACTCCCTGCCCCGACCAATACCCCTGCCAGAT 1200  
Db 1141 GCACCTGTCTGGACACAGTGTAGCACTCCCTGCCCCGACCAATACCCCTGCCAGAT 1200  
Qy 1201 GGCCGCGCGCCACTGCCCCCTGGTGGCAGTCCCTCACTGCTGCTCAATGGCTCCCTC 1260  
Db 1201 GGCCGCGCGCCACTGCCCCCTGGTGGCAGTCCCTCACTGCTGCTCAATGGCTCCCTC 1260  
Qy 1261 CACATCAGCAGCAGCTTGGTGGCCATGCAATCAGGGGTCCCTGGCCCGCGGCCCTGTGGC 1320  
Db 1261 CACATCAGCAGCAGCTTGGTGGCCATGCAATCAGGGGTCCCTGGCCCGCGGCCCTGTGGC 1320  
Qy 1321 TGCTGTCCAGCTGCCCTGAACATTTGGGAGCAAAAGAAAGTCTCTCTACTGCTGTAGGCC 1380  
Db 1321 TGCTGTCCAGCTGCCCTGAACATTTGGGAGCAAAAGAAAGTCTCTCTACTGCTGTAGGCC 1380  
Qy 1381 ACACCTCTCAAAAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTCTATGGTACCAGTT 1440  
Db 1381 ACACCTCTCAAAAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTCTATGGTACCAGTT 1440  
Qy 1441 AAAGCCAGTGGCAGCTCCCACTGCGGGGAGGCGAGCTGGAGTCTGTCTCATATATGTA 1500  
Db 1441 AAAGCCAGTGGCAGCTCCCACTGCGGGGAGGCGAGCTGGAGTCTGTCTCATATATGTA 1500

Qy 1501 TTCTTGGAGCCTGTGCCTATGTCACTGCAATTTTCAGAGACATATTTCAACTCCTCTGCT 1560  
Db 1501 TTCTTGGAGCCTGTGCCTATGTCACTGCAATTTTCAGAGACATATTTCAACTCCTCTGCT 1560  
Qy 1561 CTTTCAAAACCTTGGTGTCTATCCGGCAGAGGAGGAGGAGCAAGTGGAGCAGGCGTT 1620  
Db 1561 CTTTCAAAACCTTGGTGTCTATCCGGCAGAGGAGGAGGAGCAAGTGGAGCAGGCGTT 1620  
Qy 1621 AGCAGGAGCAGTTTCTTGCCAGAGACAGCCTGCTGCCAGCGGGGAGCCCTCATAG 1680  
Db 1621 AGCAGGAGCAGTTTCTTGCCAGAGACAGCCTGCTGCCAGCGGGGAGCCCTCATAG 1680  
Qy 1681 GAGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740  
Db 1681 GAGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740  
Qy 1741 GGTGCCCCACAGCTTCCAGGCTCTCCCTGACCTGCTCTATGCCCCACACCTACGCTG 1800  
Db 1741 GGTGCCCCACAGCTTCCAGGCTCTCCCTGACCTGCTCTATGCCCCACACCTACGCTG 1800  
Qy 1801 CCGTGGCTCTGTGAGTGTACGTAGATAGTCTGCGCTGGTCTGCTGTTTGTCTGTA 1860  
Db 1801 CCGTGGCTCTGTGAGTGTACGTAGATAGTCTGCGCTGGTCTGCTGTTTGTCTGTA 1860  
Qy 1861 AAAGCTTAATGGCTGGCCAGGCTGTGTACCTTCTCCAAAGCAAGCATATGAGGATC 1920  
Db 1861 AAAGCTTAATGGCTGGCCAGGCTGTGTACCTTCTCCAAAGCAAGCATATGAGGATC 1920  
Qy 1921 TACCCAGACTCCACCTCTGACACACACTCCTCCACCTCTCAAGCCTCCAACTCTTGGC 1980  
Db 1921 TACCCAGACTCCACCTCTGACACACACTCCTCCACCTCTCAAGCCTCCAACTCTTGGC 1980  
Qy 1981 CAGATTGGGCTCATTAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
Db 1981 CAGATTGGGCTCATTAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
Qy 2041 TGGTGGTCTGCTGTGAGCTCTTCAAGTCTTAACTTAACTTAACTTAACTTAACTTAACT 2100  
Db 2041 TGGTGGTCTGCTGTGAGCTCTTCAAGTCTTAACTTAACTTAACTTAACTTAACTTAACT 2100  
Qy 2101 TGGCTGAGACCCAGCAGCAGCTTCTGCGCTTCTCCCTGCTCAATCTTAAAGCAGT 2160  
Db 2101 TGGCTGAGACCCAGCAGCAGCTTCTGCGCTTCTCCCTGCTCAATCTTAAAGCAGT 2160  
Qy 2161 GCCACACCTCCAAAGTGGAAATAGAAAGAGTTTCAAGTAAAGGCTGCAAGGAATCTT 2220  
Db 2161 GCCACACCTCCAAAGTGGAAATAGAAAGAGTTTCAAGTAAAGGCTGCAAGGAATCTT 2220  
Qy 2221 ATCTTGGCCACATGTCTCCGTGCACACACCAATGGAGTTAACTTGGAGTTGACTAT 2280  
Db 2221 ATCTTGGCCACATGTCTCCGTGCACACACCAATGGAGTTAACTTGGAGTTGACTAT 2280  
Qy 2281 TTTAATGTCTGCCAGAGTTTAACTTCCCTCTGCTTCCCTTCTCTCTCTCTCTCTCTCT 2340  
Db 2281 TTTAATGTCTGCCAGAGTTTAACTTCCCTCTGCTTCCCTTCTCTCTCTCTCTCTCTCT 2340  
Qy 2341 AGCACACATCTTCTGCTTCCCGAGTTTCTGCGCTTCCACCTTCCAGCTTCAATGCTC 2400  
Db 2341 AGCACACATCTTCTGCTTCCCGAGTTTCTGCGCTTCCACCTTCCAGCTTCAATGCTC 2400  
Qy 2401 AGTGTGTGCTTAAATAAATGAGACATATTTTCTCTAAAAA 2447  
Db 2401 AGTGTGTGCTTAAATAAATGAGACATATTTTCTCTAAAAA 2447

RESULT 3

AK095713

LOCUS

DEFINITION

ACCESSION

VERSION

AK095713 Homo sapiens cDNA FLJ38394 fis, clone FEBRA2007534, moderately  
similar to CamK1-like protein kinase.  
AK095713  
AK095713.1 GI:21755035

2612 bp mRNA linear

PRI 15-JUL-2002

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
Source  
Query Match  
Best Local Similarity 87.6%; Score 2143; DB 9; Length 2612;  
Matches 2393; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 39 GGCATCCTCAGAGCTTCAACTCTGGAGGCAATGGTGCAGGAAAGGAGATGACTGCAG 98  
DB 215 GGCATCCTCAGAGCTTCAACTCTGGAGGCAATGGTGCAGGAAAGGAGATGACTGCAG 274  
QY 99 TTCCTGGAAGAACACACACCAACATCCGGAACCTTCATTTTATGGAAGTGCCTGGG 158  
DB 275 TTCCTGGAAGAACACACACCAACATCCGGAACCTTCATTTTATGGAAGTGCCTGGG 334  
QY 159 ATCAGGAGCTTCTCAGAAAGTTTCTGGTGAAGCAAGACTGACTGGGAAGCTCTTTGC 218  
DB 335 ATCAGGAGCTTCTCAGAAAGTTTCTGGTGAAGCAAGACTGACTGGGAAGCTCTTTGC 394  
QY 219 TCTGGAAGTGCATCAAGAGTCACTTCCTCCGGACACGACCTGGAGATGAGATTGC 278  
DB 395 TCTGGAAGTGCATCAAGAGTCACTTCCTCCGGACACGACCTGGAGATGAGATTGC 454  
QY 279 TGTGTTGAAAAGATCAAGCATGAAACATTTGTGACCTTGGAGGACATCTATGAGAGAC 338  
DB 455 TGTGTTGAAAAGATCAAGCATGAAACATTTGTGACCTTGGAGGACATCTATGAGAGAC 514  
QY 339 CACCCACTACTACTGGTCATGAGCTTGTCTGGTGGGAGCTCTTTGACCGGATCCT 398  
DB 515 CACCCACTACTACTGGTCATGAGCTTGTCTGGTGGGAGCTCTTTGACCGGATCCT 574  
QY 399 GGAGCGGGTGCTACACAGAGAGGATCCAGTCTGGTGATCCAGAGGCTTGTGCGGC 458  
|||||

oligo capping; fls (full insert sequence).  
Homo sapiens fetal brain cDNA to mRNA, clone\_lib:FEBRA2  
clone:FEBRA2007534.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hirao,M.,  
Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K.,  
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,  
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,  
Kimura,K., Yanashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,  
Takahashi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,  
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,  
Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2612)  
Isogai,T. and Yamamoto,J.  
Direct Submission  
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: HRI and  
RAB; annotation: HRI and RAB.  
Location/Qualifiers  
1. .2612  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="FEBRA2007534"  
/tissue\_type="brain"  
/clone\_lib="FEBRA2"  
/dev\_stage="fetal"  
/note="cloning vector: pME18SFL3"  
BASE COUNT 619 a 750 c 645 g 598 t

QY	1539	AGACATATTCAACTCTCTGCTCTTCCAAACCTGGTGTCTATCCGGCAGAGGGAAGG	1598
Db	1715	AGACATATTCAACTCTCTGCTCTTCCAAACCTGGTGTCTATCCGGCAGAGGGAAGG	1774
QY	1599	CAGAGCAAGTGGAGCAGGCTTAGCAGGAGCAGTTCTGGCCAGAGCACCAGCTGCTG	1658
Db	1775	CAGAGCAAGTGGAGCAGGCTTAGCAGGAGCAGTTCTGGCCAGAGCACCAGCCGCTG	1834
QY	1659	CCAGCGGGCAGCCCTCATAGAGGCCCAGGAGGAGCCCCAAGCGGTAGAACCTTGT	1718
Db	1835	CCAGCGGGCAGCCCTCATAGAGGCCCAGGAGGAGCCCCAAGCGGTAGAACCTTGT	1894
QY	1719	TGAAGCTGTGACGAGAGAGCGGTGCCACCAGCTTCCAGTCTCCCTGACCTGCC	1778
Db	1895	TGAAGCTGTGACGAGAGAGCGGTGCCACCAGCTTCCAGTCTCCCTGACCTGCC	1954
QY	1779	TCATATGCCACACCCCTAGCTGCGGTCTGTGTGAGTGTACGTAGATAGCTTCGGCT	1838
Db	1955	TCATATGCCACACCCCTAGCTGCGGTCTGTGTGAGTGTACGTAGATAGCTTCGGCT	2014
QY	1839	GGTCTGTGCTGTTTCTGTGAAAGCTTAATGGGCTGGCCAGGCTGTGTCACTTCTCC	1898
Db	2015	GGTCTGTGCTGTTTCTGTGAAAGCTTAATGGGCTGGCCAGGCTGTGTCACTTCTCC	2074
QY	1899	AAGCAAGGCCATATGAGCATCTACCCAGACTCCCACTCTGCACACACTCACTCCCACT	1958
Db	2075	AAGCAAGGCCATATGAGCATCTACCCAGACTCCCACTCTGCACACACTCACTCCCACT	2134
QY	1959	CTCAAGCCTCCAACTCTTGGCCAGATTGGGCTCATTAATGTGTTGGCTGCCCATCTGC	2018
Db	2135	CTCAAGCCTCCAACTCTTGGCCAGATTGGGCTCATTAATGTGTTGGCTGCCCATCTGC	2194
QY	2019	ATGAATGACAGGAGCTCCCATGTTGGTCTGCCTGTGAGCTTCTCAAGTTCTAATCCTT	2078
Db	2195	ATGAATGACAGGAGCTCCCATGTTGGTCTGCCTGTGAGCTTCTCAAGTTCTAATCCTT	2254
QY	2079	AACCTCAGGATTAGCTCCCAAGTGCCTGAGACCCAGCCAGCACACTTCTGCGCCCTTCTC	2138
Db	2255	AACCTCAGGATTAGCTCCCAAGTGCCTGAGACCCAGCCAGCACACTTCTGCGCCCTTCTC	2314
QY	2139	CTGCGCTCAATCTAAAGCAGTGCCACACCCCTCCAAAGTGGAAATGAAAGAGTTCATGA	2198
Db	2315	CTGCGCTCAATCTAAAGCAGTGCCACACCCCTCCAAAGTGGAAATGAAAGAGTTCATGA	2374
QY	2199	GTAAAGGCTCAAGGAATCTTATCTGCGCAGCATCTCTCGTGACACACCCCAATGA	2258
Db	2375	GTAAAGGCTCAAGGAATCTTATCTGCGCAGCATCTCTCGTGACACACCCCAATGA	2434
QY	2259	GTTAACTTTGGAAAGTTGACTATTTAATGTCTGCCAGGAGTTCTAATCCTGCTCTGTTCT	2318
Db	2435	GTTAACTTTGGAAAGTTGACTATTTAATGTCTGCCAGGAGTTCTAATCCTGCTCTGTTCT	2494
QY	2319	CTTTTCTCTCCTTGAAGTCCAGCACCATCTTGTCTTCCCTCCAGTTTCTCGCCCT	2378
Db	2495	CTTTTCTCTCCTTGAAGTCCAGCACCATCTTGTCTTCCCTCCAGTTTCTCGCCCT	2554
QY	2379	CCACCCCTCCAGCTTCATGCTCAGTGTGTGCTTAATAAATGGACATATTTTCTCT	2436
Db	2555	CCACCCCTCCAGCTTCATGCTCAGTGTGTGCTTAATAAATGGACATATTTTCTCT	2612
RESULT	4		
AF428261			
LOCUS		2464 bp	linear
DEFINITION		AF428261 Homo sapiens calcium/calmodulin-dependent protein kinase I gamma	PRI 07-NOV-2001
ACCESSION		AF428261 (CAWK1G) mRNA, complete cds.	
VERSION		AF428261.1	GI:16755791
KEYWORDS			
SOURCE		Homo sapiens	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
AUTHORS	1 (bases 1 to 2464)	
TITLE	Schutte,B.C., Bjork,B.C., Coppage,K.B., Malik,M.I., Gregory,S.G., Scott,D.J., Brentzell,I.M., Watanabe,Y., Dixon,M.J. and Murray,J.C. A preliminary gene map for the Van der Woude syndrome critical region derived from 900 kb of genomic sequence at 1q32-q41	
JOURNAL	Genome Res. 10 (1), 81-94 (2000)	
MEDLINE	20113118	
PUBMED	10645953	
REFERENCE	2 (bases 1 to 2464)	
AUTHORS	Bjork,B.C., Watanabe,Y., Murray,J.C. and Schutte,B.C.	
TITLE	Characterization of the human ortholog of rat Cam Kinase I gamma (CamK1g) at 1q32-q41	
JOURNAL	Unpublished	
REFERENCE	3 (bases 1 to 2464)	
AUTHORS	Bjork,B.C., Watanabe,Y., Murray,J.C. and Schutte,B.C.	
TITLE	Direct Submission	
JOURNAL	Submitted (05-OCT-2001) Genetics, University of Iowa, 140 EMRB, Iowa City, IA 52242, USA	
FEATURES	Location/Qualifiers	
source	1..2464	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="1"	
	/map="1q32-q41"	
gene	1..2464	
	/gene="CAWK1G"	
CDS	73..1503	
	/gene="CAWK1G"	
	/codon_start=1	
	/product="calcium/calmodulin-dependent protein kinase I gamma"	
	/protein_id="AAL28100.1"	
	/db_xref="GI:16755792"	
	/translation="MGRKEDDCSWKKQTTNIRKTFIFNEVLGSGAFSEFLVKQRL TGLFLAKCIKKYSPAFRSDSLENIETAVLKKIKHENIVTLEDIYESTHYLLVMQLVSG GELFDRLIERGVYTEDKASLVITQOVLAVKYLHENGIVHRDLKPENLLYLTPEENSKI MTFDFGLSKMEQNGINSTACGTPGYVAPELAQKPYSKAVDCWSIGVTIYLLCGYPP FYBETESKLFEKIKGYEYFESPWDDIDESAQDFICHULEKDPNERYTCERKALSHPW IDGNTALHRDIYPSVSLQIKNFPAKSKRWQAFNAAVVHMRKLMHNLHSPGIRPEVE NRPPEQASSETSRPSPETITIEAPVLDHVALPALTOIPCQHGRRPTAPGGRSLNCL VNGSLHSSLSLVPWHOGSLAAGPCGCCSCLNIGNKGSYSCSEPTLLKANKKQNFKE SEVMVPVKASGSGSHCRAGGTGVCLIN"	
polyA_signal	2415..2420	
	/gene="CAWK1G"	
BASE COUNT	609 a	708 c 602 g 545 t
ORIGIN		
Query Match	83.7%; Score 2048; DB 9; Length 2464;	
Best Local Similarity	99.8%; Pred. No. 0;	
Matches 2418; Conservative	0; Mismatches	5; Indels 1; Gaps 1;
QY	24	TCATTCCCGAGTCCCTGGCATCCTCAGAAAGCTTCAACTCTGGAGGCAATGGTGCAGGA 83
Db	27	TCATTCCCGAGTCCCTGGCATCCTCAGAAAGCTTCAACTCTGGAGGCAATGGTGCAGGA 86
QY	84	AGAAGATGACTGCAGTTCCTGGAGAAACAGACACCAACATCCGGAACACCTTCATTT 143
Db	87	AGAAGATGACTGCAGTTCCTGGAGAAACAGACACCAACATCCGGAACACCTTCATTT 146
QY	144	TATGGAAGTCTGGATCAGGAGCTTTCAGAAAGTTTTCCTGGTGAAGCAAAAGACTGAC 203
Db	147	TATGGAAGTCTGGATCAGGAGCTTTCAGAAAGTTTTCCTGGTGAAGCAAAAGACTGAC 206
QY	204	TGGGAAGCTCTTTCCTGAAAGTGCATCAAGAGTCACTACCTTCCTCGGGACAGCAGCT 263
Db	207	TGGGAAGCTCTTTCCTGAAAGTGCATCAAGAGTCACTACCTTCCTCGGGACAGCAGCT 266
QY	264	GGAGATGAGATTCTGTTCTCAAAAGATCAACCATGAAACATTTGTGACCTCGGAGGA 323
Db	267	GGAGATGAGATTCTGTTCTCAAAAGATCAACCATGAAACATTTGTGACCTCGGAGGA 326
QY	324	CATCTATGAGACACCCACCACTACTACCTGGTCTATGAGCTTTGTTCTGTTGGGAGCT 383

|||||  
Db 327 CATCTATGAGAGACACACCCACTACTACCTGGTCAATGCAGCTTGTCTGGTGGGAGCT 386  
QY 384 CTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAGAGATGCCAGTGTGGTGATCCA 443  
Db 387 CTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAGAGATGCCAGTGTGGTGATCCA 446  
QY 444 GCAGGTCTTGTGGCAGTGAATACCTACATGAGAATGGCATGCTCCACAGAGACTTAAA 503  
Db 447 GCAGGTCTTGTGGCAGTGAATACCTACATGAGAATGGCATGCTCCACAGAGACTTAAA 506  
QY 504 GCCCGAAAACCTGGCTTTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCACTGACTT 563  
Db 507 GCCCGAAAACCTGGCTTTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCACTGACTT 566  
QY 564 TGGTCTGTCCAAAGTGAAGTGAAGTGGCATCATGTCCACTGCCCTGGGACCCGAGCTA 623  
Db 567 TGGTCTGTCCAAAGTGAAGTGAAGTGGCATCATGTCCACTGCCCTGGGACCCGAGCTA 626  
QY 624 CGTGGCTCCAGAAGTGTCTGGCCAGAAACCCCTACAGCAAGGCTGTGGATTTGCTGGTCCAT 683  
Db 627 CGTGGCTCCAGAAGTGTCTGGCCAGAAACCCCTACAGCAAGGCTGTGGATTTGCTGGTCCAT 686  
QY 684 CGGGCTCATCACTACATATTGCTCTGTGGATACCCCGCTTCTATGAAGAAACGGAGTC 743  
Db 687 CGGGCTCATCACTACATATTGCTCTGTGGATACCCCGCTTCTATGAAGAAACGGAGTC 746  
QY 744 TAAGCTTTTCGAGAAGATCAAGAGGGCTACTATGAGTTTGAGTCTCCATTTCTGGGATGA 803  
Db 747 TAAGCTTTTCGAGAAGATCAAGAGGGCTACTATGAGTTTGAGTCTCCATTTCTGGGATGA 806  
QY 804 CATTTCTGATCAGCACCAAGGACTTTATTTGGCATACCCCGCTTCTATGAAGAAACGGAGTC 863  
Db 807 CATTTCTGATCAGCACCAAGGACTTTATTTGGCATACCCCGCTTCTATGAAGAAACGGAGTC 866  
QY 864 GTACACCTGTGAGAAGGCTTGTGAGTCATCCCTGGATTGACGGAACACAGGCCCTCCACCG 923  
Db 867 GTACACCTGTGAGAAGGCTTGTGAGTCATCCCTGGATTGACGGAACACAGGCCCTCCACCG 926  
QY 924 GGNATCTACCCATCAGTCAGCTCCAGTCCAGTCCAGAGAACTTTGCTAGAGCAAGTGGAG 983  
Db 927 GGNATCTACCCATCAGTCAGCTCCAGTCCAGTCCAGAGAACTTTGCTAAGAGCAAGTGGAG 986  
QY 984 GCAAGCTTTCAACGACGACGCTGTGGTGCACACATGAGGAAGCTTACACATGAACCTGCA 1043  
Db 987 GCAAGCTTTCAACGACGACGCTGTGGTGCACACATGAGGAAGCTTACACATGAACCTGCA 1046  
QY 1044 CAGCCGCGGCTCCGCCAGAGGTGGAGACAGGCCGCTGAAACTCAAGCTCAGAAAC 1103  
Db 1047 CAGCCGCGGCTCCGCCAGAGGTGGAGAACAGGCCGCTGAAACTCAAGCTCAGAAAC 1106  
QY 1104 CTCTAGACCCAGCTCCCTGAGATCACCATCACCAGGACACTGTCTCTGGACACAGTGT 1163  
Db 1107 CTCTAGACCCAGCTCCCTGAGATCACCATCACCAGGACACTGTCTCTGGACACAGTGT 1166  
QY 1164 AGCACTCCCTGGCCGTGACCCAAATACCTTGCCAGCATGGCCGCGGCCCACTGCCCTGG 1223  
Db 1167 AGCACTCCCTGGCCGTGACCCAAATACCTTGCCAGCATGGCCGCGGCCCACTGCCCTGG 1226  
QY 1224 TGGCAGTCCCTCAACTGCTGTGTCATATGGCTCCCTCCACATCAGCAGCAGCTGGTGCC 1283  
Db 1227 TGGCAGTCCCTCAACTGCTGTGTCATATGGCTCCCTCCACATCAGCAGCAGCTGGTGCC 1286  
QY 1284 CATGTCATCAGGGTCCCTGGCCGCGGCCCTGTGGCTGCTGCTCCAGCTGCCTGAACAT 1343  
Db 1287 CATGTCATCAGGGTCCCTGGCCGCGGCCCTGTGGCTGCTGCTCCAGCTGCCTGAACAT 1346  
QY 1344 TGGGAGCAAGGAAAGTCTCTACTGTCTGTGAGCCCACTCTCAAAAAGGCCAACAA 1403  
Db 1347 TGGGAGCAAGGAAAGTCTCTACTGTCTGTGAGCCCACTCTCAAAAAGGCCAACAA 1406  
QY 1404 AAACAGAACTTCAAGTCGGAGTCAATGTTACAGTTAAAGCCAGTGGCAGCTCCCACTG 1463  
|||||

RESULT 5  
HS2721161

Db 1407 AAACAGAACTTCAAGTCGGAGGTCAATGTTACCAAGTTAAAGCCAGTGGCAGCTCCCACTG 1466  
QY 1464 CCGGCGAGGCGAGCTGGAGTCTGTCTCATTTATGTGATTTCTGGAGCTGTGCCTATGTCT 1523  
Db 1467 CCGGCGAGGCGAGCTGGAGTCTGTCTCATTTATGTGATTTCTGGAGCTGTGCCTATGTCT 1526  
QY 1524 ACTGCAATTTTCAGGAGACATATCAACTCTCTCTCTTCCAAACCTGTGTCTATCCG 1583  
Db 1527 ACTGCAATTTTCAGGAGACATATCAACTCTCTCTCTTCCAAACCTGTGTCTATCCG 1586  
QY 1584 GCAGAGGAGAGAGCAGAGCAAGTGGAGCGGCTTAGCAGGAGCAGTTTCTGGCCAGA 1643  
Db 1587 GCAGAGGAGAGAGCAGAGCAAGTGGAGCGGCTTAGCAGGAGCAGTTTCTGGCCAGA 1646  
QY 1644 AGCACAGCCTCTGCCAGCGGCGAGCCCTCATAGAGGCCCCAGGAGGAGGCCCAAG 1703  
Db 1647 AGCACAGCCTCTGCCAGCGGCGAGCCCTCATAGAGGCCCCAGGAGGAGGCCCAAG 1706  
QY 1704 GGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGGTCT 1763  
Db 1707 GGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGGTCT 1766  
QY 1764 CCCTGACCTGCTCTATGCCCCACACCCCTACCTGACGCTGGCTGTGTGCAAGTGTAGCT 1823  
Db 1767 CCCTGACCTGCTCTATGCCCCACA - CCTACGTGCGGTGCTCTGTGCAAGTGTAGCT 1825  
QY 1824 AGATAGCTCTCGCTGGGTCTGTGCTGTTGCTGTGCTGAAAAGCTTAATGGGTGGCCAGGC 1883  
Db 1826 AGATAGCTCTCGCTGGGTCTGTGCTGTTGCTGTGAAAAGCTTAATGGGTGGCCAGGC 1885  
QY 1884 TGTGTCACTTCTCCAAGCAAAAGCCATATGGAGCATCTACCCAGACTCCCACTCTGCACA 1943  
Db 1886 TGTGTCACTTCTCCAAGCAAAAGCCATATGGAGCATCTACCCAGACTCCCACTCTGCACA 1945  
QY 1944 CACTCACTCCCACCTCTCAAGCCTCCAACTCTTGGCAGATTTGGGCTCAATTAATGTGCT 2003  
Db 1946 CACTCACTCCCACCTCTCAAGCCTCCAACTCTTGGCAGATTTGGGCTCAATTAATGTGCT 2005  
QY 2004 TGCTGCCCATCTGCATGAATGACAGGAGCTCCCACTGTGTGCTGTGAGTCTCTT 2063  
Db 2006 TGCTGCCCATCTGCATGAATGACAGGAGCTCCCACTGTGTGCTGTGAGTCTCTT 2065  
QY 2064 CAAGTCTTAATCCTTAACCTCCAGGATTAAGTCTCCCAAGTGGCTGAGACCCAGCCAGCA 2123  
Db 2066 CAAGTCTTAATCCTTAACCTCCAGGATTAAGTCTCCCAAGTGGCTGAGACCCAGCCAGCA 2125  
QY 2124 CTTCTGGCCCTTCTCCCTGCCTCAATCTAAAAGCAGTGCACACCTTCCAAAGTGAATA 2183  
Db 2126 CTTCTGGCCCTTCTCCCTGCCTCAATCTAAAAGCAGTGCACACCTTCCAAAGTGAATA 2185  
QY 2184 GAAAGAGTTTCATGATAAGGCTGCAAGGAAATTTCTATCTTGGCCACATGTCTCCCTGG 2243  
Db 2186 GAAAGAGTTTCATGATAAGGCTGCAAGGAAATTTCTATCTTGGCCACATGTCTCCCTGG 2245  
QY 2244 CACACACCAATGGAGTTAACTTGAAGTTCGACATTTTAAATGCTTGCAGGAGTTCTA 2303  
Db 2246 CACACACCAATGGAGTTAACTTGAAGTTCGACATTTTAAATGCTTGCAGGAGTTCTA 2305  
QY 2304 ATCTGCTCTGTTCCTCTTCTCTCTTGAAGTTCAGCAGACCACTTCTTGTCTCTCC 2363  
Db 2306 ATCTGCTCTGTTCCTCTTCTCTCTTGAAGTTCAGCAGACCACTTCTTGTCTCTCC 2365  
QY 2364 CAGTTCTCTGGCCCTCCACCCCTCCAGCTTTCATGTGCTGTGTGCTTAAATAATGGA 2423  
Db 2366 CAGTTCTCTGGCCCTCCACCCCTCCAGCTTTCATGTGCTGTGTGCTTAAATAATGGA 2425  
QY 2424 CATATTTTCTCTAAAAA 2447  
Db 2426 CATATTTTCTCTAAAAA 2449  
|||||

LOCUS HS272L161 1738 bp mRNA linear PRI 21-APR-1999  
DEFINITION Human gene isolated from PAC 272L16, chromosome 1, similar to  
calcium/calmodulin dependent protein kinases.  
ACCESSION AL049688  
VERSION AL049688.1 GI:4678721  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1738)  
Rhodes,S.  
Direct Submission  
Submitted (21-APR-1999) E-mail contact: humquery@sanger.ac.uk  
This sequence was generated from cDNA clones isolated using  
sequence from the bacterial clone 272L16 (AL023754) and EST data.  
The EST sequences listed match this sequence with an identity of at  
least 95% between the coordinates shown.  
Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr1/ Partial, experimentally  
determined gene.  
Sanger Centre name: dJ272L16.Cl.1.  
Location/Qualifiers  
1..1738  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="lq32.1-32.3"  
/contig="1448"  
/codon\_start=3  
/product="hypothetical protein"  
/protein\_id="CAB41259.1"  
/db\_xref="GI:4678722"  
/translation="STLEAMGRKEEDDCSWKKOTTNIRKTFIFMEVLGSAFSEVFL  
VKORLTGKLFALKCIKKSPAPRDSLENAVLKKIKHENIVLTEDIYESTTHVYLVM  
QVSGGGLDFRLRGVYTEKDSLVIQOVL SAVKYLHENGIVHRDLKPNLXLITPE  
ENSKIMITDFGLSMEQNGINSTAGTGPVAVPEVLQAQKYSKAVDCWSIGVITYILL  
CGYPFYEETSEKLFKEIKEGYEFESFPWDDISESAKDFICHLLEKDPNERVCEKA  
LSHPWIDNTALHRDIYPSVSLQIQKFAKSKWRQAFNAAVVHMKRLHMLHSPGV  
RPEVNERPETQASETSRPSPEITTEAPLDHVALPALPOLPCQHGRRPTAPGGR  
SLNCLVNGSLHISSLVPMHOGSLAAGPCGCCSCLNIGSKGSSYCSSEPTLLKANK  
KONFKSEVMVPVKASGSSHCRAGQTGVCLIM"  
<1..107  
/number=1  
108..236  
/number=2  
237..311  
/number=3  
312..450  
/number=4  
451..574  
/number=5  
575..650  
/number=6  
651..763  
/number=7  
653..777  
/note="matches EST AI215131 from clone IMAGE:1925595"  
764..842  
/number=8  
843..930  
/number=9  
868..1159  
/note="matches EST AA351937"  
931..1355  
/number=10  
1356..1483  
/number=11  
join(1469..1550,1553..1723,1721..1738)  
/note="matches EST R05661 from clone 29500"  
1484..1736  
/number=12

CDS

FEATURES  
source

ORIGIN

Query Match 66.9%; Score 1636; DB 9; Length 1738;  
Best Local Similarity 99.9%; Pred No. 0;  
Matches 1736; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 53 CTTCAACTCTGGAGGCAATGGGTGCGAAAGGAAGAAGATGACTGCACTTCTTGAAGTTCCTTGAAGAAC 112  
Db 1 CTTCAACTCTGGAGGCAATGGGTGCGAAAGGAAGAAGATGACTGCACTTCTTGAAGTTCCTTGAAGAAC 60  
QY 113 AGACCACCAACATCCCGAAACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTTCT 172  
Db 61 AGACCACCAACATCCCGAAACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTTCT 120  
QY 173 CAGAAGTTTCTGGTGAAGCAAGACTGACTGGGAAGCTCTTTTGTCTCTGAAGTGCATCA 232  
Db 121 CAGAAGTTTCTGGTGAAGCAAGACTGACTGGGAAGCTCTTTTGTCTCTGAAGTGCATCA 180  
QY 233 ASAAGTCACTGCTTCCGGGACAGCAGCTGAGGAATGAGATTTGCTGTGTTGAAAGA 292  
Db 181 AGAAGTCACTGCTTCCGGGACAGCAGCTGAGGAATGAGATTTGCTGTGTTGAAAGA 240  
QY 293 TCAAGCATGAAACATTTGACCCCTGGAGACATCTATGAGACACACCCACTACTACC 352  
Db 241 TCAAGCATGAAACATTTGACCCCTGGAGACATCTATGAGACACACCCACTACTACC 300  
QY 353 TGGTCATGACAGCTTGTCTGGTGGGAGCTCTTTGACCGGATCCTTGGAGCGGGTGTCT 412  
Db 301 TGGTCATGACAGCTTGTCTGGTGGGAGCTCTTTGACCGGATCCTTGGAGCGGGTGTCT 360  
QY 413 ACACAGAGAAGATGCGAGTCTGGTATCAGCAGGCTTTGCGGAGTGAATACCTAC 472  
Db 361 ACACAGAGAAGATGCGAGTCTGGTATCAGCAGGCTTTGCGGAGTGAATACCTAC 420  
QY 473 ATGAGATGCATCTGCCACAGAGACTTAAAGCCCGAAACCTGTCTTACCTTACCCCTG 532  
Db 421 ATGAGATGCATCTGCCACAGAGACTTAAAGCCCGAAACCTGTCTTACCTTACCCCTG 480  
QY 533 AAGAGAACTCTAAGATCATGATCAGTCTTGGTCTGCTCAAGATGGAACAGATGGCA 592  
Db 481 AAGAGAACTCTAAGATCATGATCAGTCTTGGTCTGCTCAAGATGGAACAGATGGCA 540  
QY 593 TCATGTCCACTGCTGTGGACCCAGGCTACGTGCTCCAGAGTGTGGCCCAAGAAC 652  
Db 541 TCATGTCCACTGCTGTGGACCCAGGCTACGTGCTCCAGAGTGTGGCCCAAGAAC 600  
QY 653 CCTACAGCAAGGCTGTGGATTGCTGGTCCATCGGCGTCACTACATATTGCTCTGTG 712  
Db 601 CCTACAGCAAGGCTGTGGATTGCTGGTCCATCGGCGTCACTACATATTGCTCTGTG 660  
QY 713 GATACCCCGGTTCTATGAAGAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGGGGCT 772  
Db 661 GATACCCCGGTTCTATGAAGAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGGGGCT 720  
QY 773 ACTATGAGTTTGTAGTCTCCATCTTGGATGACATTTCTGAGTCAGCAAGAGCTTTATT 832  
Db 721 ACTATGAGTTTGTAGTCTCCATCTTGGATGACATTTCTGAGTCAGCAAGAGCTTTATT 780  
QY 833 GCCACTTTGCTTGAAGAGGATCCGAACAGCGGTACACCTGTGGAAGAGGCTTGTAGTCATC 892  
Db 781 GCCACTTTGCTTGAAGAGGATCCGAACAGCGGTACACCTGTGGAAGAGGCTTGTAGTCATC 840  
QY 893 CTTGGATTGACGGAACACAGCCCTCCACGGGACATCTACCATCAGTCAGCCTCCAGA 952  
Db 841 CTTGGATTGACGGAACACAGCCCTCCACGGGACATCTACCATCAGTCAGCCTCCAGA 900  
QY 953 TCCAGAGAAGCTTTGCTTAAGAGCAAGTGGAGGCAAGCTTCAACGACGAGCTGTGGTGC 1012  
Db 901 TCCAGAGAAGCTTTGCTTAAGAGCAAGTGGAGGCAAGCTTCAACGACGAGCTGTGGTGC 960  
QY 1013 ACCACATGAGGAAGCTTACATGAACTGACAGAGCCGGGCTCCGCCAGAGGTTGGAGA 1072  
Db 961 ACCACATGAGGAAGCTTACATGAACTGACAGAGCCGGGCTCCGCCAGAGGTTGGAGA 1020

BASE COUNT 432 a 491 c 458 g 357 t

QY	1073	ACAGGCGCCTGAACCTCAAGCCTCAGAAACCTCTAGACCCAGCTCCCTCGATCACCA	1132
Db	1021	ACAGGCGCCTGAACCTCAAGCCTCAGAAACCTCTAGACCCAGCTCCCTCGATCACCA	1080
QY	1133	TCACCGAGGACCTGCTCGGACGACAGTGTAGCACTCCCTGCCCTGACCCAAATTACCT	1192
Db	1081	TCACCGAGGACCTGCTCGGACGACAGTGTAGCACTCCCTGCCCTGACCCAAATTACCT	1140
QY	1193	GCCAGCATGCCCGCGCCACTGCCCTCGTGGGAGGTCCTCAACTGCCCTGGTGGCAATG	1252
Db	1141	GCCAGCATGCCCGCGCCACTGCCCTCGTGGGAGGTCCTCAACTGCCCTGGTGGCAATG	1200
QY	1253	GCTCCCTCCACATCAGCAGAGCCTGGTCCCATGATCAGGGGTCCCTGGCGCGCGGC	1312
Db	1201	GCTCCCTCCACATCAGCAGAGCCTGGTCCCATGATCAGGGGTCCCTGGCGCGCGGC	1260
QY	1313	CCTGTGCTGCTGCTCCAGCTGCTCAACATTTGGGAGCAAGAAAGTCTCTACTGCT	1372
Db	1261	CCTGTGCTGCTGCTCCAGCTGCTCAACATTTGGGAGCAAGAAAGTCTCTACTGCT	1320
QY	1373	CTGAGCCACATCTCTCAAAAGGCCAACAAAAACAGAACTTCAAGTGGAGGTCATGG	1432
Db	1321	CTGAGCCACATCTCTCAAAAGGCCAACAAAAACAGAACTTCAAGTGGAGGTCATGG	1380
QY	1433	TACCAGTTAAGCCAGTGGCAGCTCCACTGCGGGCAGGCGAGACTGGAGTCTGTCTCA	1492
Db	1381	TACCAGTTAAGCCAGTGGCAGCTCCACTGCGGGCAGGCGAGACTGGAGTCTGTCTCA	1440
QY	1493	TTATGTGATTCCTGGAGCCTGTGCTATGCTCACTGCAATTTTCAGGAGACATATTCACT	1552
Db	1441	TTATGTGATTCCTGGAGCCTGTGCTATGCTCACTGCAATTTTCAGGAGACATATTCACT	1500
QY	1553	CCTGTGCTCTTCAACCTGGTGTCTATCCGGCAGAGGAGGAAGCAGAGCAAGTGGAG	1612
Db	1501	CCTGTGCTCTTCAACCTGGTGTCTATCCGGCAGAGGAGGAAGCAGAGCAAGTGGAG	1560
QY	1613	CAGGCTTTAGCAGAGCAGTTCTGCGCAGAGACCCAGCTGCTGCCAGCGGCGAGCC	1672
Db	1561	CAGGCTTTAGCAGAGCAGTTCTGCGCAGAGACCCAGCTGCTGCCAGCGGCGAGCC	1620
QY	1673	CCTCATAGGAGGCCAGGAGGAGCCCAAGGCGTAGAGCCCTGTTGAAGCTGTGAGCA	1732
Db	1621	CCTCATAGGAGGCCAGGAGGAGCCCAAGGCGTAGAGCCCTGTTGAAGCTGTGAGCA	1680
QY	1733	GGAGAACGGTGGCCACCAGCTTCAGGCTCCTGACCTGCTGCTCTATGCCCCAC	1790
Db	1681	GGAGAACGGTGGCCACCAGCTTCAGGCTCCTGACCTGCTGCTCTATGCCCCAC	1738
RESULT	6	HS272L16	157875 bp DNA linear PRI 23-NOV-1999
LOCUS	HS272L16		Human DNA sequence from clone 272L16 on chromosome 1q32.1-32.3.
DEFINITION			Contains the 3' end of the LAMB3 gene for Laminin, Beta 3 (Nicein, Kalinin, BM600) and a novel Rat Ca2+/Calmodulin dependent Protein Kinase LIKE gene. Contains ESTs, STSS, GSSs, genomic marker D1S491 and a ca repeat polymorphism, complete sequence.
ACCESSION	AL023754		
VERSION	AL023754.1	GI:4007152	
KEYWORDS	HTG; BM600; ca repeat polymorphism; Ca2+/Calmodulin dependent Protein Kinase; D1S491; Kalinin; LAMB3; Laminin Beta 3; Nicein.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 157875)		
TITLE	Grafham,D.		
JOURNAL	Direct Submission		
	Submitted (27-NOV-1998) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk		
COMMENT	On Dec 12, 1998 this sequence version replaced gi:3873472.		

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence is the entire insert of clone 272L16. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

272L16 is from the library RPCI1 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:pcrPAC2> This sequence was generated from a human chromosome 1 bacterial clone contig constructed in collaboration by the Sanger Centre chromosome 1 mapping group and Brian Schutte, Bryan Bjork, Kevin Coppage and Jeffrey Murray. Department of Pediatrics, University of Iowa, USA. Further information can be found at <http://www.sanger.ac.uk/HGP/Chrl>.

FEATURES	Location/Qualifiers
source	1..157875
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="1"
	/map="q32.1-32.3"
	/clone="RPI-272L16"
	/clone_lib="RPCI-1"
repeat_region	966..1078
	/note="MLT1I repeat: matches 281..410 of consensus"
repeat_region	1082..1506
	/note="MSRD repeat: matches 1..426 of consensus"
repeat_region	1585..1875
	/note="MLT1A1 repeat: matches 1..319 of consensus"
repeat_region	1898..2024
	/note="MLT1I repeat: matches 58..203 of consensus"
repeat_region	2347..2925
	/note="MER34 repeat: matches 6..543 of consensus"
repeat_region	3090..3417
	/note="MLT1A2 repeat: matches 23..374 of consensus"
repeat_region	3677..4074
	/note="MER54B repeat: matches 486..902 of consensus"
repeat_region	4089..5338
	/note="MLT1A2-internal repeat: matches 358..1643 of consensus"
repeat_region	5339..5532
	/note="AluJo repeat: matches 101..302 of consensus"
repeat_region	5536..5966
	/note="MLT1A1-internal repeat: matches 5..450 of consensus"
repeat_region	5972..6019
	/note="MLT1A1 repeat: matches 318..365 of consensus"
repeat_region	6020..6459
	/note="LRR7 repeat: matches 1..450 of consensus"
repeat_region	6460..6781
	/note="MLT1A1 repeat: matches 1..318 of consensus"
repeat_region	6786..6870
	/note="MIR repeat: matches 56..142 of consensus"
repeat_region	8020..8067
	/note="MIR repeat: matches 95..146 of consensus"
repeat_region	8150..8293
	/note="MIR repeat: matches 68..212 of consensus"
repeat_region	8380..9083
	/note="MER50 repeat: matches 1..711 of consensus"
repeat_region	9084..9303
	/note="MIR repeat: matches 13..262 of consensus"
repeat_region	9493..9795
	/note="Alusx repeat: matches 1..304 of consensus"
repeat_region	10108..10282
	/note="L2 repeat: matches 2506..2695 of consensus"
repeat_region	10929..11005

repeat_region	/note="L2 repeat: matches 2671. .2748 of consensus" 12092. .12129	repeat_region	29821. .30076 /note="HAL1 repeat: matches 355. .618 of consensus"
repeat_region	/note="19 copies 2 mer ca 84% conserved" 12095. .12130	repeat_region	30095. .30228 /note="FLAW_A repeat: matches 7. .142 of consensus"
repeat_region	/note="9 copies 4 mer acac 86% conserved" 12722. .12855	repeat_region	30246. .30440 /note="HAL1 repeat: matches 683. .874 of consensus"
repeat_region	/note="MER91A repeat: matches 20. .186 of consensus" 13108. .13221	repeat_region	30539. .30788 /note="Charliela repeat: matches 1189. .1455 of consensus"
misc_feature	/note="MIR repeat: matches 80. .191 of consensus" complement(13563. .13979)	repeat_region	30789. .31099 /note="Alusg repeat: matches 1. .306 of consensus"
repeat_region	/note="match: GSSs AQ075457 AQ076848" 14235. .14279	repeat_region	31100. .31167 /note="Charliela repeat: matches 1125. .1189 of consensus"
repeat_region	/note="MIR repeat: matches 60. .105 of consensus" 14327. .14442	repeat_region	31170. .31914 /note="L1M2 repeat: matches 5448. .6190 of consensus"
repeat_region	/note="MIR repeat: matches 112. .249 of consensus" 14495. .15046	repeat_region	31923. .32141 /note="L1M2 repeat: matches 4317. .4530 of consensus"
repeat_region	/note="L2 repeat: matches 2127. .2750 of consensus" 16671. .16859	repeat_region	32142. .32551 /note="MSTA repeat: matches 5. .426 of consensus"
repeat_region	/note="MER5A repeat: matches 2. .189 of consensus" 17103. .17476	repeat_region	32552. .33436 /note="L1M2 repeat: matches 3451. .4317 of consensus"
repeat_region	/note="L2 repeat: matches 17. .422 of consensus" 17679. .17938	repeat_region	33437. .33756 /note="Aluy repeat: matches 1. .309 of consensus"
repeat_region	/note="MIR repeat: matches 7. .262 of consensus" 17891. .17953	repeat_region	33757. .36041 /note="L1M2 repeat: matches 761. .3451 of consensus"
repeat_region	/note="L2 repeat: matches 2694. .2749 of consensus" 18650. .18758	repeat_region	37817. .38850 /note="Charliela repeat: matches 1. .1142 of consensus"
repeat_region	/note="L2 repeat: matches 2641. .2750 of consensus" 19246. .19406	repeat_region	39293. .39437 /note="MIR repeat: matches 1. .140 of consensus"
repeat_region	/note="MER5A repeat: matches 4. .189 of consensus" 19460. .19871	misc_feature	complement(39554. .39839)
repeat_region	/note="L2 repeat: matches 2309. .2748 of consensus" 19916. .19987	repeat_region	/note="match: GSS AQ070531" 41165. .41445
repeat_region	/note="MIR repeat: matches 63. .140 of consensus" 20200. .20228	repeat_region	/note="L2 repeat: matches 2180. .2489 of consensus" 41615. .41810
repeat_region	/note="MER34 repeat: matches 510. .538 of consensus" 20233. .20365		
repeat_region	/note="MER34 repeat: matches 413. .543 of consensus" 20338. .20388		
repeat_region	/note="LNR29 repeat: matches 454. .503 of consensus" 20459. .20809		
repeat_region	/note="MER47A repeat: matches 2. .366 of consensus" 20976. .21143		
repeat_region	/note="MER34 repeat: matches 6. .172 of consensus" 21834. .22278		
repeat_region	/note="MLT1C repeat: matches 1. .466 of consensus" 22477. .22528		
repeat_region	/note="26 copies 2 mer ag 79% conserved" 22826. .22937		
repeat_region	/note="MER81 repeat: matches 1. .112 of consensus" 24897. .25034		
repeat_region	/note="MER5A repeat: matches 30. .171 of consensus" 25084. .25193		
repeat_region	/note="MIR repeat: matches 26. .145 of consensus" 25207. .25435		
repeat_region	/note="MIR repeat: matches 13. .259 of consensus" 25919. .26363		
repeat_region	/note="MLT1F repeat: matches 68. .541 of consensus" 26364. .26813		
repeat_region	/note="MER4A2 repeat: matches 1. .503 of consensus" 26814. .26885		
repeat_region	/note="MLT1F repeat: matches 1. .68 of consensus" 26931. .27357		
repeat_region	/note="L2 repeat: matches 2292. .2723 of consensus" 27779. .27820		
repeat_region	/note="21 copies 2 mer ca 100% conserved" 27779. .27818		
repeat_region	/note="10 copies 4 mer caca 100% conserved" 29036. .29195		
repeat_region	/note="HAL1 repeat: matches 4. .163 of consensus" 29263. .29500		
repeat_region	/note="HAL1 repeat: matches 141. .355 of consensus" 29501. .29820		
repeat_region	/note="MLT1A1 repeat: matches 1. .365 of consensus"		

Query Match 36.9% Score 903; DB 9; Length 157875;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1535	CAGGAGACATATCAACTCCTCTCTTCCAAACCTGGTGTCTATCCGGCAGAGGAGG	1594	
Db	147496	CAGGAGACATATCAACTCCTCTCTTCCAAACCTGGTGTCTATCCGGCAGAGGAGG	147555	
Qy	1595	AAGCAGAGCAAGTGGAGCGGCTTAGCAGGACGCTTTCTGCCAGAGCACCAGCCT	1654	
Db	147556	AAGCAGAGCAAGTGGAGCGGCTTAGCAGGACGCTTTCTGCCAGAGCACCAGCCT	147615	
Qy	1655	GCTGCCAGCGGGCAGCGCCCTCATAGGAGGCCAGGAGGCCAGGCCAGGCGTAGAGCC	1714	
Db	147616	GCTGCCAGCGGGCAGCGCCCTCATAGGAGGCCAGGAGGCCAGGCCAGGCGTAGAGCC	147675	
Qy	1715	TGTTGAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGGTCTCCCTGACCTGC	1774	
Db	147676	TGTTGAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGGTCTCCCTGACCTGC	147735	
Qy	1775	CTGCTCTATGCCACACACCCCTACGTCCGCTGTGTGCAGTGTACGTAGATAGCTCTC	1834	
Db	147736	CTGCTCTATGCCACACACCCCTACGTCCGCTGTGTGCAGTGTACGTAGATAGCTCTC	147795	
Qy	1835	GCCTGGGTCTGTGCTTTGTCTGTAAGGCTTAATGGGCTGGCCAGGCTGTGTACCTT	1894	
Db	147796	GCCTGGGTCTGTGCTTTGTCTGTAAGGCTTAATGGGCTGGCCAGGCTGTGTACCTT	147855	
Qy	1895	CTCCAAAGCAAGCCATATGGAGCATCTACCCAGACTCCCACTTCGCACACTCACCCTCC	1954	
Db	147856	CTCCAAAGCAAGCCATATGGAGCATCTACCCAGACTCCCACTTCGCACACTCACCCTCC	147915	
Qy	1955	ACCTCTCAAGCCTCCAACCTCTTGGCCAGATTGGGCTCATTAATGTCTGCTGCCCTGC	2014	
Db	147916	ACCTCTCAAGCCTCCAACCTCTTGGCCAGATTGGGCTCATTAATGTCTGCTGCCCTGC	147975	
Qy	2015	CTGCATGAATGACAGGCACTCCCATGGTGTGCTGTGAGCTCTTCAAGTTCTAAT	2074	
Db	147976	CTGCATGAATGACAGGCACTCCCATGGTGTGCTGTGAGCTCTTCAAGTTCTAAT	148035	

QY 2075 CCTTAAGTCCAGATTAGCTCCCAAGTGGCTGAGACCCAGCCAGCACATCTTGGCCCT 2134  
Db 148036 CCTTAAGTCCAGATTAGCTCCCAAGTGGCTGAGACCCAGCCAGCACATCTTGGCCCT 148095  
QY 2135 TCTCCCTGCTCAATCTAAAGCAGTGCACACCCCTCCAAAGTGAATAGAAAGATTC 2194  
Db 148096 TCTCCCTGCTCAATCTAAAGCAGTGCACACCCCTCCAAAGTGAATAGAAAGATTC 148155  
QY 2195 ATGAGTAAGGGTGAAGGAATCTTATCTCGGCCACATGCTCCCTCGTGCACACCCCAA 2254  
Db 148156 ATGAGTAAGGGTGAAGGAATCTTATCTCGGCCACATGCTCCCTCGTGCACACCCCAA 148215  
QY 2255 TGGAGTAACTTGAAGTGAATCTTATCTCGGCCACATGCTCCCTCGTGCACACCCCAA 2314  
Db 148216 TGGAGTAACTTGAAGTGAATCTTATCTCGGCCACATGCTCCCTCGTGCACACCCCAA 148275  
QY 2315 GTTCCCTTTCT 2374  
Db 148276 GTTCCCTTTCT 148335  
QY 2375 CCTCCACCCCTCCAGCTTCATGCTAGTGTGCTTAAATAAATGACATATTTTCT 2434  
Db 148336 CCTCCACCCCTCCAGCTTCATGCTAGTGTGCTTAAATAAATGACATATTTTCT 148395  
QY 2435 CTA 2437  
Db 148396 CTA 148398

RESULT 7  
HUMD1S70R/c  
LOCUS HUMD1S70R 1249 bp DNA linear PRI 17-MAR-1994  
DEFINITION Human dinucleotide repeat polymorphism at the D1S70 locus.  
ACCESSION L26254  
VERSION L26254.1 GI:460666  
KEYWORDS dinucleotide repeat polymorphism.  
SOURCE Homo sapiens DNA.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1249)  
AUTHORS Weston,M.D. and Kimberling,W.J.  
TITLE A new polymorphism for the D1S70 locus  
JOURNAL Unpublished (1993)  
FEATURES  
source  
1..1249  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="1q41"  
748..771  
/standard\_name="dinucleotide repeat polymorphism"  
/note="putative"  
/rpt\_type=tandem  
BASE COUNT 310 a 229 c 391 g 319 t  
ORIGIN  
Query Match 3.2%; Score 78; DB 9; Length 1249;  
Best Local Similarity 100.0%; Pred. No. 7.7e-33;  
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 628 GCTCCAGAGTCTGGCCAGAAACCCCTACAGCAAGGCTGTGGATTGCTGTGTCATCGGC 687  
Db 987 GCTCCAGAGTCTGGCCAGAAACCCCTACAGCAAGGCTGTGGATTGCTGTGTCATCGGC 928  
QY 688 GTCATCACTACATATG 705  
Db 927 GTCATCACTACATATG 910  
RESULT 8  
BC021840  
LOCUS BC021840 2416 bp mRNA linear ROD 07-AUG-2002

DEFINITION Mus musculus, clone MGC:30513 IMAGE:4502479, mRNA, complete cds.  
ACCESSION BC021840  
VERSION BC021840.1 GI:18256866  
KEYWORDS MGC.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 2416)  
AUTHORS Strausberg,R.  
TITLE Direct Submission  
JOURNAL Submitted (18-JAN-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
REMARK NIH-MGC project URL: http://mgc.nci.nih.gov  
COMMENT Contact: MGC help desk  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
Contact: amg@bcm.tmc.edu  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,  
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAC Plate: 41 Row: i Column: 13  
This clone was selected for full length sequencing because it  
passed the following selection criteria: GenomesScan gene  
prediction.  
FEATURES  
source  
1..2416  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="MGC:30513 IMAGE:4502479"  
/tissue\_type="Eye, retina, mouse strain C57BL/6"  
/clone\_lib="NIH\_MGC\_94"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
87..1520  
/codon\_start=1  
/product="Unknown (protein for MGC:30513)"  
/protein\_id="AAH21840.1"  
/db\_xref="GI:18256867"  
/translation="MGRKEEDCSWKKQTNIKTFIFMEVLGSGAFSEFLVKQRV  
TGKFLAIKSPAFRDSLEIAVLKRIKHENIVTDIYEYTHYLVMLVSG  
GELFDRIILRGVTEKDALVIQVLSAVKYLHENGIVHRDLKPNLILYLPENSKI  
MITDFGLSKMEQNGVMSTACGTPGVAPVLAOKPYSKAVDCWSTGVITYILLGCTPP  
FYEETESKLFEKIKFEGYEFESFPWDDIISAKDFICHLEKDPNERVTCFKALHPW  
IDGNTALHREDIYPSVSLQIKNFASKWRQAFNAAVYVHHMKRLHMLHLSPSVQREVE  
NRPPVSPAPEVSDHSDSITEAPILDPPTLPALTRPCSHSRSPSPAPSGGRLNC  
LVNGSLRISSSLVPMQOGLATPGCGCCSCLNTGNGKGSYCSEPTLFRKANKQNF  
KSEVMVPYKAGSTHCRGGQTGCVLM"  
BASE COUNT 572 a 694 c 602 g 548 t  
ORIGIN  
Query Match 2.9%; Score 71; DB 10; Length 2416;  
Best Local Similarity 100.0%; Pred. No. 9.2e-29;  
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 304 AACATTGTGACCTGGAGGACATCTATGAGACACACCCACTACTACCTGTCATGCAG 363  
Db 321 AACATTGTGACCTGGAGGACATCTATGAGACACACCCACTACTACCTGTCATGCAG 380  
QY 364 CTGTTTCTCTGG 374  
Db 381 CTGTTTCTCTGG 391

```
RESULT 9
AF428262
LOCUS
DEFINITION AF428262 2427 bp mRNA linear ROD 07-NOV-2001
Mus musculus calcium/calmodulin-dependent protein kinase I gamma
(Camk1g) mRNA, complete cds.
ACCESSION AF428262
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 2427)
AUTHORS Bjork.B.C., Watanabe.Y., Murray.J.C. and Schutte.B.C.
TITLE Characterization of the human ortholog of rat Cam Kinase I gamma
(CamK1g) at Iq32-q41
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 2427)
AUTHORS Bjork.B.C., Watanabe.Y., Murray.J.C. and Schutte.B.C.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-2001) Genetics, University of Iowa, 140 EMBR,
Iowa City, IA 52242, USA
FEATURES
source
Location/Qualifiers
1..2427
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="1"
/map="1H, 104.00 cM"
/clone="IMAGE:4502479"
/tissue_type="retina"
/note="corresponds to EST BG293660"
1..2427
/gene="Camk1g"
86..1519
/gene="Camk1g"
/codon_start=1
/product="calcium/calmodulin-dependent protein kinase I
gamma"
/protein_id="AAL28101.1"
/db_xref="GI:16755794"
/translation="MGRKEEDCSWKKQTNIKRTIFMEVLGSGAFSEVLVKQRY
TKGLFALCAIKSPAFRDSLEIAVLKRIKHENIVTLEDIYESTHYLVWQLVSG
GELFDRLERGVYTKDASLVIOQVLSAVKYLHENGIVHRDLKPENLLYITPENSKI
MITDFGLSKMEQNGVMSTACGTPGYVAPVLAQKPYSKAVDCWSIGVITYILICGYPP
PYETESKLFEKIKEGYEFESPFWDIDISAKDFICHLLEKDPNERYTCEKALRHPW
IDGNTALRDLYPVSLOIQKNEAKSKWROAFNAAVVHMRKLMHNLHSPSVROEVE
NRPVSPAPEVSRDSDSHSITEAPILDPTPLPALJRLPCSHSRSPSPSGRSLNC
LYNGSLRISLSLVPWQOGPLATPGCGCCSCINIGNKGSYCISEPTLFRANKQNF
KSEVMVPVKAGGSTHCRGGQTGVCLVM"
BASE COUNT 584 a 694 c 601 g 547 t 1 others
ORIGIN
Query Match 2.9%; Score 71; DB 10; Length 2427;
Best Local Similarity 100.0%; Pred. No. 9.2e-29;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 304 AACATTGTGACCTGGAGGACATCTATGAGACGACACCCACTACTACTGTCATGCAG 363
|||||
Db 320 AACATTGTGACCTGGAGGACATCTATGAGACGACACCCACTACTACTGTCATGCAG 379
|||||
QY 364 CTGTTCTCGG 374
|||||
Db 380 CTGTTCTCGG 390
|||||
RESULT 10
AL365314
LOCUS
DEFINITION AL365314 182054 bp DNA linear ROD 29-JUN-2002
Mouse DNA sequence from clone RP23-106N23 on chromosome 1, complete
sequence.
ACCESSION AL365314
VERSION
KEYWORDS
SOURCE
Mus musculus.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 182054)
AUTHORS Kay.M.
TITLE Direct Submission
JOURNAL Submitted (29-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jun 13, 2001 this sequence version replaced gi:13396595.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-106N23 is
from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6
-----
Center: Genome Center
Center: UK Medical Research Council
Center code: UK-MRC
Web site: http://mrcseq.har.mrc.ac.uk
Contact: mouseq@har.mrc.ac.uk
-----
FEATURES
source
Location/Qualifiers
1..182054
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="1"
/clone="RP23-106N23"
/clone_lib="RPCI-23"
BASE COUNT 49155 a 44402 c 42300 g 46197 t
ORIGIN
Query Match 2.5%; Score 62; DB 10; Length 182054;
Best Local Similarity 100.0%; Pred. No. 1.7e-23;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 304 AACATTGTGACCTGGAGGACATCTATGAGACGACACCCACTACTACTGTCATGCAG 363
|||||
Db 84769 AACATTGTGACCTGGAGGACATCTATGAGACGACACCCACTACTACTGTCATGCAG 84828
|||||
QY 364 CT 365
||
Db 84829 CT 84830
|||||
RESULT 11
AC022675/c
LOCUS
DEFINITION AC022675 184754 bp DNA linear HTG 09-MAY-2001
Mus musculus clone RP23-59N15, WORKING DRAFT SEQUENCE, 13 unordered
pieces.
ACCESSION AC022675
VERSION AC022675.3 GI:13940729
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus.
```

## ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 184754)  
Birren.B., Linton.L., Nusbaum.C. and Lander.E.  
Unpublished  
2 (bases 1 to 184754)  
Birren.B., Linton.L., Nusbaum.C., Lander.E., Abraham.H., Allen.N.,  
Anderson.S., Baldwin.J., Barna.N., Beckerly.R., Beda.F.,  
Boguslavskiy.L., Boukhgalter.B., Brown.A., Burkett.G., Castle.A.,  
Choepe.Y., Colangelo.M., Collins.S., Collymore.A., Cooke.P.,  
DeArellano.K., Dewar.K., Domino.M., Doyle.M., Fenestor.J.,  
Ferreira.P., FitzHugh.W., Forrest.C., Gage.D., Galagan.J.,  
Gardyna.S., Grant.G., Hagos.B., Heaford.A., Horton.L.,  
Howland.J.C., Johnson.R., Jones.C., Kann.L., Karatas.A., Klein.J.,  
Lander.S., Lehoczy.J., Levine.R., Lieu.C., Liu.G., Locke.K.,  
Macdonald.P., Marquis.N., McEwan.P., McGurk.A., McKernan.K.,  
McPheeters.R., Meldrim.J., Meneus.L., Morrow.J., Naylor.J.,  
Norman.C.H., O'Connor.T., O'Donnell.P., Olivari.T.M., Peterson.K.,  
Pierre.N., Pisani.C., Pollara.V., Raymond.C., Riley.R., Rothman.D.,  
Roy.A., Santos.R., Severy.P., Spencer.B., Stange-Thomann.N.,  
Stojanovic.N., Subramanian.A., Talamas.J., Tesfaye.S., Theodore.J.,  
Tirrell.A., Vassiliev.H., Viel.R., Vo.A., Wu.X., Wyman.D., Ye.W.J.,  
Zimmer.A. and Zody.M.

TITLE  
JOURNAL

Direct Submission  
Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 4, 2001 this sequence version replaced gi:7657723.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

## COMMENT

-----  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
-----  
Project Information  
Center project name: L5039  
Center clone name: 59\_N15  
-----  
Summary Statistics  
Sequencing vector: M13; M77815; 22% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 179592 bases at least Q40  
Consensus quality: 181970 bases at least Q30  
Insert size: 182000; agarose-fp  
Insert size: 183554; sum-of-contigs  
Quality coverage: 12.9 in Q20 bases; agarose-fp  
Quality coverage: 12.8 in Q20.

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 13 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 4354: contig of 4354 bp in length  
\* 4355 4454: gap of 100 bp  
\* 4455 5467: contig of 1013 bp in length  
\* 5468 5567: gap of 100 bp  
\* 5568 6717: contig of 1150 bp in length  
\* 6718 6817: gap of 100 bp  
\* 6818 8112: contig of 1295 bp in length  
\* 8113 8212: gap of 100 bp  
\* 8213 9269: contig of 1057 bp in length  
\* 9270 9369: gap of 100 bp  
\* 9370 11162: contig of 1793 bp in length  
\* 11163 11262: gap of 100 bp  
\* 11263 12591: contig of 1319 bp in length

\* 12582 12681: gap of 100 bp  
\* 12682 13892: contig of 1211 bp in length  
\* 13893 13992: gap of 100 bp  
\* 13993 23297: contig of 9305 bp in length  
\* 23298 23397: gap of 100 bp  
\* 23398 38920: contig of 15523 bp in length  
\* 38921 39020: gap of 100 bp  
\* 39021 57984: contig of 18964 bp in length  
\* 57985 58084: gap of 100 bp  
\* 58085 141207: contig of 83123 bp in length  
\* 141208 141307: gap of 100 bp  
\* 141308 184754: contig of 43447 bp in length.

## FEATURES

Location/Qualifiers  
1..184754  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone\_lib="RP23-59N15"  
/clone\_lib="RP23-59N15" female Mouse BAC"

## misc\_feature

1..4354  
/note="assembly\_fragment"

clone\_end:SP6  
vector\_side:left"

## misc\_feature

4455..5467  
/note="assembly\_fragment"

## misc\_feature

5568..6717  
/note="assembly\_fragment"

## misc\_feature

6818..8112  
/note="assembly\_fragment"

## misc\_feature

8213..9269  
/note="assembly\_fragment"

## misc\_feature

9370..11162  
/note="assembly\_fragment"

## misc\_feature

11263..12581  
/note="assembly\_fragment"

## misc\_feature

12682..13892  
/note="assembly\_fragment"

## misc\_feature

13993..23297  
/note="assembly\_fragment"

## misc\_feature

23398..38920  
/note="assembly\_fragment"

## misc\_feature

39021..57984  
/note="assembly\_fragment"

## misc\_feature

58085..141207  
/note="assembly\_fragment"

## misc\_feature

141308..184754  
/note="assembly\_fragment"

clone\_end:T7  
vector\_side:right"

BASE COUNT 47717 a 42686 c 43285 g 49746 t 1320 others  
ORIGIN

Query Match 2.5%; Score 62; DB 2; Length 184754;

Best Local Similarity 100.0%; Pred.No.1.7e-23; Length 184754;  
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 AACATTGTGACCTGGAGGACATCTATGAGAGCACCACCCTACTACTGCTCATGCAG 363  
|||||  
Db 140281 AACATTGTGACCTGGAGGACATCTATGAGAGCACCACCCTACTACTGCTCATGCAG 140222

QY 364 CT 365

Db 140221 CT 140220

## RESULT 12

D86557 D86557 1013 bp mRNA linear ROD 07-FEB-1999  
LOCUS Rattus norvegicus mRNA for Protein Kinase, partial cds.  
DEFINITION D86557  
ACCESSION D86557  
VERSION D86557.1 GI:2077933  
KEYWORDS Rattus norvegicus embryo (E18) brain cDNA to mRNA, clone\_lib:s.  
SOURCE Nakanishi clone:N5.

```

ORGANISM      Rattus norvegicus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.
REFERENCE
AUTHORS      Yokokura,H., Terada,O., Naito,Y. and Hidaka,H.
TITLE        Isolation and comparison of rat cDNAs encoding
              Ca2+/calmodulin-dependent protein kinase I isoforms
JOURNAL      Biochim. Biophys. Acta 1338 (1), 8-12 (1997)
MEDLINE      97228532
REFERENCE
AUTHORS      Yokokura,H.
TITLE        Direct Submission
JOURNAL      Submitted (15-JUL-1996) Hisayuki Yokokura, Nagoya University School
              of Medicine, Department of Pharmacology; Tsurumai 65, Showa-ku,
              Nagoya, Aichi 466, Japan (Tel:052-744-2075, Fax:052-744-2083)
FEATURES
source
1. .1013
   /organism="Rattus norvegicus"
   /db_xref="taxon:10116"
   /clone="N5"
   /tissue_type="brain"
   /clone_lib="S. Nakanishi"
   /dev_stage="embryo (E18)"
   <1. .86
87. >1013
   /note="similar to calcium/calmodulin-dependent protein
       kinase I mRNA (L24907, L26288)"
   /codon_start=1
   /product="Protein Kinase"
   /protein_id="BAAL19880.1"
   /db_xref="GI:2077934"
   /translation="MGRKEEDCSWKQKTTNIRKTFIEVLGSGAFSEFLVKQRV
       TGLFALCKIKSPAFRDSSENEIIVLKRKHENIVTLEDIYESTHYLYMQLVSG
       GELFRILRGVYTPKDSLVISAVKYLHENGIVHRDLKPENLLYLPPENSKI
       MTFDEGLSKMQNGYMGACGTGPGVAPEVLAQRPYSKAVCSGIVITYLILGYP
       FYEETFLKFKIKGYEFESPFWDIDISEAKDFICHLLKDPNERYTCCKALRHFW
       IDGNFALHRDIPVSVLQIQNFAKSKWRQAFN"
BASE COUNT   276 a 245 c 262 g 230 t
ORIGIN
Query Match      2.4%; Score 59; DB 10; Length 1013;
Best Local Similarity 100.0%; Pred. No. 8.6e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 778 GAGTTTGAGTCTCATTCTGGGATGACATTTCTGAGTCAGCCAGACGATTTATTGGCCA 836
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 795 GAGTTTGAGTCTCATTCTGGGATGACATTTCTGAGTCAGCCAGACGATTTATTGGCCA 853

RESULT 13
AC126166
LOCUS        AC126166
DEFINITION  Rattus norvegicus clone CH230-ID9, *** SEQUENCING IN PROGRESS ***,
              76 unordered pieces.
ACCESSION   AC126166
VERSION     AC126166.2 GI:21700365
KEYWORDS    HTG; HTGS; PHASE1.
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.
REFERENCE
AUTHORS      Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
              Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
              Barbara,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
              Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
              Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
              Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
              Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
              Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
              Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,

```

```

Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Woodson,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 178272)
Worley,K.C.
Direct Submission
Submitted (04-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 178272)
Worley,K.C.
Direct Submission
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 6, 2002 this sequence version replaced gi:21693754.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: TUCQ
Center clone name: CH230-ID9
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 113604 bases at least Q40
Consensus quality: 121900 bases at least Q30
Consensus quality: 127034 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
      (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
      consists of 76 contigs. The true order of the pieces
      is not known and their order in this sequence record is
      arbitrary. Gaps between the contigs are represented as
      runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
      as soon as it is available and the accession number will
      be preserved.
      1
      1071: contig of 1071 bp in length

```

```
* 1072 1171: gap of unknown length
* 1172 2571: contig of 1400 bp in length
* 2572 2671: gap of unknown length
* 2672 4193: contig of 1522 bp in length
* 4194 4293: gap of unknown length
* 4294 5757: contig of 1464 bp in length
* 5758 5857: gap of unknown length
* 5858 6960: contig of 1103 bp in length
* 6961 7061: gap of unknown length
* 7062 8110: contig of 1050 bp in length
* 8111 9302: contig of 1092 bp in length
* 9303 9402: gap of unknown length
* 9403 10661: contig of 1259 bp in length
* 10662 10761: gap of unknown length
* 10762 12254: contig of 1493 bp in length
* 12255 12354: gap of unknown length
* 12355 13368: contig of 1014 bp in length
* 13369 13468: gap of unknown length
* 13469 14799: contig of 1331 bp in length
* 14800 14899: gap of unknown length
* 14900 15912: contig of 1013 bp in length
* 15913 16012: gap of unknown length
* 16013 17129: contig of 1117 bp in length
* 17130 17229: gap of unknown length
* 17230 18571: contig of 1342 bp in length
* 18572 18671: gap of unknown length
* 18672 20220: contig of 1449 bp in length
* 20221 20220: gap of unknown length
* 20221 21583: contig of 1363 bp in length
* 21584 21683: gap of unknown length
* 21684 23016: contig of 1333 bp in length
* 23017 23116: gap of unknown length
* 23117 24283: contig of 1167 bp in length
* 24284 24383: gap of unknown length
* 24384 25500: contig of 1117 bp in length
* 25501 25501: gap of unknown length
* 25501 27309: contig of 1709 bp in length
* 27310 27409: gap of unknown length
* 27410 29157: contig of 1748 bp in length
* 29158 29257: gap of unknown length
* 29258 31559: contig of 2302 bp in length
* 31560 31659: gap of unknown length
* 31660 32799: contig of 1140 bp in length
* 32800 32899: gap of unknown length
* 32900 35290: contig of 2391 bp in length
* 35291 35390: gap of unknown length
* 35391 36925: contig of 1535 bp in length
* 36926 37025: gap of unknown length
* 37026 38876: contig of 1851 bp in length
* 38877 38976: gap of unknown length
* 38977 40184: contig of 1208 bp in length
* 40185 40284: gap of unknown length
* 40285 41512: contig of 1228 bp in length
* 41513 41612: gap of unknown length
* 41613 42937: contig of 1325 bp in length
* 42938 43037: gap of unknown length
* 43038 44247: contig of 1210 bp in length
* 44248 44347: gap of unknown length
* 44348 46092: contig of 1745 bp in length
* 46093 46192: gap of unknown length
* 46193 47274: contig of 1082 bp in length
* 47275 47374: gap of unknown length
* 47375 48878: contig of 1504 bp in length
* 48879 48979: gap of unknown length
* 48979 50315: contig of 1337 bp in length
* 50316 50415: gap of unknown length
* 50416 52545: contig of 2130 bp in length
* 52546 52645: gap of unknown length
* 52646 54617: contig of 1972 bp in length
* 54618 54717: gap of unknown length
* 54718 55940: contig of 1223 bp in length
* 55941 56040: gap of unknown length

* 56041 58002: contig of 1962 bp in length
* 58003 58102: gap of unknown length
* 58103 59810: contig of 1708 bp in length
* 59811 59910: gap of unknown length
* 59911 61918: contig of 2008 bp in length
* 61919 62018: gap of unknown length
* 62019 63896: contig of 1878 bp in length
* 63897 65703: contig of 1707 bp in length
* 65704 65803: gap of unknown length
* 65804 67650: contig of 1847 bp in length
* 67651 67750: gap of unknown length
* 67751 69467: contig of 1717 bp in length
* 69468 71839: contig of 2272 bp in length
* 71840 71939: gap of unknown length
* 71940 74071: contig of 2132 bp in length
* 74072 74171: gap of unknown length
* 74172 76853: contig of 2682 bp in length
* 76854 76953: gap of unknown length
* 76954 79190: contig of 2237 bp in length
* 79191 79290: gap of unknown length
* 79291 80794: contig of 1504 bp in length
* 80795 80894: gap of unknown length
* 80895 82989: contig of 2095 bp in length
* 82990 83089: gap of unknown length
* 83090 85392: contig of 2303 bp in length
* 85393 85492: gap of unknown length
* 85493 88314: contig of 2822 bp in length
* 88315 88414: gap of unknown length
* 88415 90655: contig of 2241 bp in length
* 90656 90755: gap of unknown length
* 90756 94227: contig of 3472 bp in length

Query Match 1.6%; Score 40; DB 2; Length 178272;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 778 GAGTTGAGTCTCCATTCTGGGATGACATTTCGTGAGTCAG 817
|||||
Db 91915 GAGTTGAGTCTCCATTCTGGGATGACATTTCGTGAGTCAG 91954

RESULT 14
AC120477/c
LOCUS AC120477 191584 bp DNA linear HTG 23-JUL-2002
DEFINITION Rattus norvegicus clone CH230-33N8, *** SEQUENCING IN PROGRESS ***,
71 unordered pieces.
ACCESSION AC120477
VERSION AC120477.3 GI:21902854
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 191584)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haviak,P., Hawes,A., Hernandez,J.,
```

Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louissege,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Mosier,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G.,  
Oragunye,N., Oviado,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,  
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Tsamani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs.R.  
Direct Submission  
Unpublished  
2 (bases 1 to 191584)  
Worley,K.C.  
Direct Submission  
Submitted (07-MAY-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 191584)  
Worley,K.C.  
Direct Submission  
Submitted (23-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 18, 2002 this sequence version replaced gi:20522179.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GXLL  
Center clone name: CH230-33N8  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 138862 bases at least Q40  
Consensus quality: 145657 bases at least Q30  
Consensus quality: 150087 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 71 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 1006: contig of 1006 bp in length  
\* 1007 1106: gap of unknown length  
\* 1107 2111: contig of 1005 bp in length  
\* 2112 2211: gap of unknown length  
\* 2212 3227: contig of 1016 bp in length  
\* 3228 3327: gap of unknown length  
\* 3328 4335: contig of 1008 bp in length  
\* 4336 4435: gap of unknown length  
\* 4436 5441: contig of 1006 bp in length  
\* 5441 5541: gap of unknown length  
\* 5542 6587: contig of 1046 bp in length  
\* 6588 6687: gap of unknown length  
\* 6688 8648: contig of 1961 bp in length  
\* 8649 9951: contig of 1203 bp in length  
\* 9952 10051: gap of unknown length  
\* 10052 11081: contig of 1030 bp in length  
\* 11082 12212: contig of 1031 bp in length  
\* 12213 13478: contig of 1166 bp in length  
\* 13479 13578: gap of unknown length  
\* 13579 14929: contig of 1351 bp in length  
\* 14930 15029: gap of unknown length  
\* 15030 16090: contig of 1061 bp in length  
\* 16091 16190: gap of unknown length  
\* 16191 17244: contig of 1054 bp in length  
\* 17245 17345: gap of unknown length  
\* 17346 18956: contig of 1612 bp in length  
\* 18957 20171: contig of 1115 bp in length  
\* 20172 20271: gap of unknown length  
\* 20272 21283: contig of 1012 bp in length  
\* 21284 21383: gap of unknown length  
\* 21384 23582: contig of 2199 bp in length  
\* 23583 23682: gap of unknown length  
\* 23683 24984: contig of 1302 bp in length  
\* 24985 25084: gap of unknown length  
\* 25085 26255: contig of 1171 bp in length  
\* 26256 27622: contig of 1267 bp in length  
\* 27623 27722: gap of unknown length  
\* 27723 29025: contig of 1303 bp in length  
\* 29026 29125: gap of unknown length  
\* 29126 30966: contig of 1841 bp in length  
\* 30967 31066: gap of unknown length  
\* 31067 32073: contig of 1007 bp in length  
\* 32074 32173: gap of unknown length  
\* 32174 33601: contig of 1428 bp in length  
\* 33602 33701: gap of unknown length  
\* 33702 34879: contig of 1178 bp in length  
\* 34880 34979: gap of unknown length  
\* 34980 36112: contig of 1133 bp in length  
\* 36113 36212: gap of unknown length  
\* 36213 38536: contig of 2324 bp in length  
\* 38537 38636: gap of unknown length  
\* 38637 40088: contig of 1452 bp in length  
\* 40089 40188: gap of unknown length  
\* 40189 43004: contig of 2816 bp in length  
\* 43005 43104: gap of unknown length  
\* 43105 44472: contig of 1368 bp in length  
\* 44473 44572: gap of unknown length  
\* 44573 47031: contig of 2459 bp in length  
\* 47032 47131: gap of unknown length  
\* 47132 49092: contig of 1961 bp in length  
\* 49093 49192: gap of unknown length  
\* 49193 51804: contig of 2612 bp in length  
\* 51805 53859: contig of 1955 bp in length  
\* 53860 53959: gap of unknown length  
\* 53960 55351: contig of 1392 bp in length  
\* 55352 57797: contig of 2346 bp in length  
\* 57798 57897: gap of unknown length  
\* 57898 60119: contig of 2222 bp in length  
\* 60120 61884: contig of 1665 bp in length  
\* 61885 61984: gap of unknown length  
\* 61985 65400: contig of 3416 bp in length  
\* 65401 65500: gap of unknown length  
\* 65501 67405: contig of 1905 bp in length

\* 67406 67505: gap of unknown length  
\* 67506 6850: contig of 1145 bp in length  
\* 6851 6850: gap of unknown length  
\* 6851 70564: contig of 1814 bp in length  
\* 70565 70564: gap of unknown length  
\* 70665 72651: contig of 1987 bp in length  
\* 72652 72751: gap of unknown length  
\* 72752 75008: contig of 2257 bp in length  
\* 75009 75108: gap of unknown length  
\* 75109 77573: contig of 2465 bp in length  
\* 77574 77674: gap of unknown length  
\* 77675 79219: contig of 1546 bp in length  
\* 79220 79319: gap of unknown length  
\* 79320 81724: contig of 2405 bp in length  
\* 81725 81824: gap of unknown length  
\* 81825 84362: contig of 2538 bp in length  
\* 84363 84462: gap of unknown length  
\* 84463 86195: contig of 1733 bp in length  
\* 86196 86295: gap of unknown length  
\* 86296 89725: contig of 3430 bp in length  
\* 89726 89825: gap of unknown length  
\* 89826 93336: contig of 3511 bp in length  
\* 93337 93437: gap of unknown length  
\* 93437 96552: contig of 3116 bp in length  
\* 96553 96552: gap of unknown length  
\* 96553 99918: contig of 3266 bp in length

Query Match 1.6%; Score 40; DB 2: Length 191584;  
Best Local Similarity 100.0%; Pred. No. 1.1e-10;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 778 GAGTTTGAGTCTCCATTCCTGGGATGACATTTCTGAGTCAG 817  
|||||  
Db 49589 GAGTTTGAGTCTCCATTCCTGGGATGACATTTCTGAGTCAG 49550

## RESULT 15

BC014825  
LOCUS BC014825 1448 bp mRNA linear ROD 07-AUG-2002  
DEFINITION Mus musculus, Similar to calcium/calmodulin-dependent protein kinase I, clone MGC:18933 IMAGE:3969343, mRNA, complete cds.  
ACCESSION BC014825  
VERSION BC014825.1 GI:15928725  
KEYWORDS MGC.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1448)  
Strausberg, R.  
Direct Submission  
Submitted (01-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

## REMARK

## COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 23 Row: j Column: 19

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

## FEATURES

## source

Location/Qualifiers

1..1448

/organism="Mus musculus"

/db\_xref="taxon:10090"

/map="FVB/N"

/clone="MGC:18933 IMAGE:3969343"

/tissue\_type="Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."

/clone\_lib="NCI CGAP\_Mam1"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

120..1244

## CDS

/codon\_start=1

/product="Similar to calcium/calmodulin-dependent protein kinase I"

/protein\_id="AAH14825.1"

/db\_xref="GI:15928726"

/translation="MPGAVEGPRWQAEIRIDYFRDLVLTGAFSEVILAEDEKRTQK  
LVAIKCIAKKALEGKSGMENEIATLHKIKHNIVALDDIYESGGLYLIMQLVSGGE  
LFDRIVERGFYTERDASRLIFQVLDVAVKYLHDLGIVHRDLKPENLLYSLDEDSKIMI  
SDPGLSKMEDPGSVLSTAGCTPGYVAPEVLAOKPYSKAVDCWSIGYAVILLCGYPPF  
YDENAKLFEQILKAEYEDSPYWDIDSDAKDFIRHLMEKDPKRFCEQALQHPWI  
AGDTALDKNIHQSVSEQIKKFAKSKWQAFNATAVVRHMRKLQLGTSQEGOGQTGSH  
GELLTPTAGGPAAGCCCRDCCVEPGESELPAPPSSRAMD"

BASE COUNT 364 a 375 c 408 g 301 t

## ORIGIN

Query Match 1.3%; Score 32; DB 10: Length 1448;

Best Local Similarity 100.0%; Pred. No. 4.4e-06;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 CCCTACAGCAAGGCTGTGGATTGCTGTCAT 683

|||||

Db 699 CCCTACAGCAAGGCTGTGGATTGCTGTCAT 730

Search completed: March 15, 2003, 04:12:48

Job time : 8188 secs



GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 15, 2003, 04:12:52 ; Search time 2223 Seconds  
(without alignments)  
3467.860 Million cell updates/sec

Title: US-09-960-643-2  
Perfect score: 2513  
Sequence: 1 MGRKEEDDCSSWKKQTTNR.....VKAGSSSHCRAGQGVCLIM 476

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09960643/runat\_07032003\_090517\_19885/app\_query.fasta\_1.647  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MAPRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09960643 @CGN.1.1.1906 @runat\_07032003\_090517\_19885 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_htc.\*  
9: gb\_estl.\*  
10: gb\_est2.\*  
11: gb\_htc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: gb\_gss.\*  
18: em\_gss\_hum.\*  
19: em\_gss\_inv.\*  
20: em\_gss\_pln.\*  
21: em\_gss\_vrt.\*  
22: em\_gss\_fun.\*  
23: em\_gss\_mam.\*  
24: em\_gss\_mus.\*  
25: em\_gss\_other.\*  
26: em\_gss\_pro.\*  
27: em\_gss\_rtd.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1532	61.0	1126	13	BM547443
2	1298	51.7	905	14	BQ934044
3	1257	50.0	809	13	BI821474
4	1234.5	49.1	812	13	BI772626
5	1191.5	47.4	1129	14	BM807335
6	1158	46.1	1051	9	AL560091
7	1147	45.6	740	13	BI824483
8	1140.5	45.4	1068	14	BM921532
9	1134	45.1	742	13	BI818261
10	1122	44.6	731	13	BI753035
11	1103	43.9	818	12	BG293660
12	1101	43.8	824	12	BG715920
13	1013.5	40.3	918	14	BQ949629
c	985	39.2	555	10	AW251224
14	965	38.5	962	9	AL556476
15	938.5	37.3	620	14	BM944418
16	920.5	36.6	573	14	BQ866330
17	901.5	35.9	1068	14	BM919261
18	894.5	35.6	873	12	BG168668
19	893.5	35.6	554	13	BI834635
20	879.5	35.0	929	13	BI758466
21	879.5	35.0	929	13	BI758466
22	856.5	34.1	871	10	BE569540
23	855.5	34.0	1052	14	BQ666820
24	849	33.8	656	13	BI667965
25	836	33.3	627	10	AW826802
26	830.5	33.0	973	13	BM051435
27	820.5	32.7	822	13	BI545545
c	813	32.4	451	10	AW254051
28	812.5	32.3	798	12	BG702279
29	800	31.8	659	13	BI463345
30	797.5	31.7	607	9	AI595095
31	797.5	31.7	817	13	BI756289
32	797.5	31.7	725	12	BG302505
33	793	31.6	636	10	BE266955
34	793	31.6	636	10	BE266955
c	790	31.4	941	9	AL578723
35	785	31.2	439	10	AW522224
36	784.5	31.2	863	13	BI757461
37	784.5	31.1	646	14	BM728430
38	780.5	30.9	832	12	BF582563
39	777.5	30.7	748	12	BF143967
40	772.5	30.5	556	12	BG895406
41	766.5	30.5	929	14	BQ897490
42	766	30.4	1061	14	BM903712
43	765	30.4	758	13	BI761267
44	763.5	30.4	758	13	BI761267
45	759.5	30.2	640	12	BG307424

# ALIGNMENTS

RESULT 1  
BM547443  
LOCUS  
DEFINITION AGENCOURT\_6507688 NIH\_MGC\_125 Homo sapiens cDNA clone IMAGE:5724450  
5', mRNA sequence.  
ACCESSION BM547443  
VERSION BM547443.1 GI:18781213  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1126)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgraphs-r@mail.nih.gov  
Tissue Procurement: Invitrogen  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM12713 row: 1 column: 19  
High quality sequence stop: 623.  
Location/Qualifiers  
1. .1126  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5724450"  
/clone\_lib="NIH\_MGC\_125"  
/lab\_host="DH10B"  
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;  
Site\_1: EcoRV (destroyed); Site\_2: NotI; RNA source pool  
of three ovaries, from females ranging in age from 38 to  
49 yo. Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 2.1 kb, insert size range 1-3.5 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 036."

BASE COUNT 299 a 291 c 286 g 244 t 6 others  
ORIGIN

## Alignment Scores:

Pred. No.: 3,44e-146 Length: 1126  
Score: 1532.00 Matches: 316  
Percent Similarity: 90.70% Conservative: 6  
Best Local Similarity: 89.01% Mismatches: 27  
Query Match: 60.96% Indels: 7  
DB: 13 Gaps: 2

US-09-960-643-2 (1-476) x BM547443 (1-1126)

Qy 1 MetGlyArgLysGluGluAspCysSerTrpLysLysGlnThrThrAsnIleArg 20  
Db 69 ATGGTGGGAAGGAGGATGACTGCTGCTGGAAGAACAGACACCATCCGG 128  
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
Db 129 AAAACCTTCATTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAGATTTTCTCGTG 188  
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60  
Db 189 AAGCAAGACTGACTGGGAAGCTTTTGTCTGAAGTGCATCAAGAGTCACCTGCCTTC 248  
Qy 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80  
Db 249 CGGACAGACCTCGGAGATGAGATGCTGCTGGAAGATCAAGCATGAACATC 308  
Qy 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100  
Db 309 GTGACCTGGAGGACATCTATGAGACCAACCCACTACTACCTGCTGTCAGCTGTGT 368  
Qy 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120  
Db 369 TCTGTGGGAGCTCTTTGACCGGATCCTTGGAGCGGGTGTCTACACAGAGAGATGCC 428  
Qy 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140  
Db 429 AGTCTGTGTATCCAGCAGGCTTGTGTGGCAGTGAATACCTACATGAGATGCGTCGTC 488  
Qy 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160  
Db 489 CACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCTGAAGAACTCTAAGATC 548

Qy 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180  
Db 549 ATGATCACTGACTTTGGTCTGTCCAAGATGAGACAGATCATGTCCACTGCGTGT 608  
Qy 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200  
Db 609 GGGACCCAGAGCTACGGTGGCTCCAGAGTGTGCGCCAGAAACCTACAGCAAGCGTGT 668  
Qy 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyr 220  
Db 669 GATTGCTGCTCATCGGCTCATCACCTACATATTGCTCTGTGGTACCCCATTTCTAT 728  
Qy 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240  
Db 729 GAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAGTTTGAGTCT 788  
Qy 241 ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHis-LeuLeuGluLys 260  
Db 789 CCATTCTGGGATGACATTCTTGAGTCAGCAAGGACTTTATTGCCACTTTGCTTGAGAA 848  
Qy 260 sAspProAsnGluArgTyrThr-CysGluLysAlaLeu---SerHisProTrpIleAspG 279  
Db 849 AGATCCGAACGAGCGGTACACCTGNTGAGAAAGGCTTGAGTCTATCCCTGGGATTGACG 908  
Qy 279 LysThrAla-LeuHisArg-AspIleTyrProSerValSerLeuGlnIleGlnLysAs 298  
Db 909 GAAACACAGCCCTCCACCGGNGACATCTACCCATCAGTCAGCTTACAGATCCAGAGAA 968  
Qy 298 nPheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetAr 318  
Db 969 CTGTGCTAANAACCAATGGAGGAGGAGGCTTCAACCCANCAAGCTGTGGTGGCCCTGAG 1028  
Qy 318 gLysLeuHisMetAsnLeuHisSerProGlyValArgProGluVal---GluAsnArgPr 337  
Db 1029 GAAAGTACC-CTGAACCTGGACAGCGGGAATTCNNCCCAAGAGGGGGGAAACAGGCC 1087  
Qy 337 oProGluThrGlnAlaSerGluThrSerArgProSer 349  
Db 1088 GCCCTGAAATCCAGCCCTCAAAACCTTTGAACCCAGT 1124

## RESULT 2

BO934044 905 bp mRNA linear EST 21-AUG-2002  
LOCUS AGENCOURT\_8802332 NIH\_MGC\_130 Mus musculus cDNA clone IMAGE:6336727  
DEFINITION 5', mRNA sequence.  
ACCESSION BO934044  
VERSION BO934044.1 GI:22349427  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 905)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgraphs-r@mail.nih.gov  
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,  
Ph.D.  
cDNA Library Preparation: ResGen, Invitrogen Corp  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM13801 row: d column: 08  
High quality sequence stop: 689.  
Location/Qualifiers  
1. .905  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6336727"

## FEATURES

source

```
/clone_lib="NIH_MGC_130"
/lab_host="DH10B (phage-resistant)"
/note="Organ: otocysts; Vector: pCMV-SPORT6.1.cdb;
Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally.
Primer: Oligo dt. Average insert size 1.95 kb.
Constructed by ResGen, Invitrogen Corp. Note: this is a
NIH_MGC Library."
```

BASE COUNT 238 a 217 c 240 g 210 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,42e-122 Length: 905  
Score: 1298.00 Matches: 256  
Percent Similarity: 98.12% Conservative: 5  
Best Local Similarity: 96.24% Mismatches: 1  
Query Match: 51.65% Indels: 4  
DB: 14 Gaps: 0

US-09-960-643-2 (1-476) x BQ934044 (1-505)

QY 1 MetGlyArgLysGluAspCysSerTrpLysLysGlnThrAsnIleArg 20  
|||||  
Db 107 ATGGGCGTAAGGAGGAGGAGTCTGCTGCTGGAAGAAACACACCAACATCAGG 166  
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
|||||  
Db 167 AAAACCTTCATCTTCATGAAGTGTGGGATCAGGAGCTTCTCAGAGGTGTCTCTGGTG 226  
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60  
|||||  
Db 227 AAGCAAAAGATGACTGGGAACTCTTGTCTGCTGAAATGATCAAGAAGTCAACGACCTTC 286  
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80  
|||||  
Db 287 CGGAGACAGACCTAGAGATGAGATCGCTGTGTGAAAGAGATCAAGCATGAGAACATTT 346  
QY 81 ValThrLeuGluAspIleTrpGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100  
|||||  
Db 347 GTGACCTGTGGAGACATCTATGAGAGCACCCACTACTACTCTGGTCATGCGAGTGTGT 406  
QY 101 SerGlyGlyLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120  
|||||  
Db 407 TCTGGAGGTGAGCTTTTACCGGATCCCTAGAGCGTGGTGTCTACACAGAAAGGATGCC 466  
QY 121 SerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140  
|||||  
Db 467 AGCCTGTGTCATCAGCAGGCTGTGTCTGCGGTGAATACCTTCATGAGAATGGCATCGTC 526  
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160  
|||||  
Db 527 CACAGAGATCTAAAGCCTGAAACCTGTGTACCTCACCCCTGAGGAGAACTCCAAGATC 586  
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180  
|||||  
Db 587 ATGATCACTGACCTTGGTCTATCCAGATGGAGATGGAGTATGTCACACGCTGT 646  
QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200  
|||||  
Db 647 GGGACCCAGGCTACGTGGCTCCAGAAGTGTGGCCAGAGCCCTACAGTAAGGCTGTG 706  
QY 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuLysGlyTyrProPheThr 220  
|||||  
Db 707 GACTGTGGTTCATTTGGTGTCAATACATACATCTGCTGTGGGTATTCCTCTTCAT 766  
QY 221 GluGluThrGluSerLysLeu-PheGluLysIleLysGluGlyTyrTyrGlu-PheGlu 240  
|||||  
Db 767 GAAGAAACAAATCAAGCTTTTTCAGAGATCAAAAGAGGCTACTACGAGGTTTGAGT 826  
QY 240 erProPheTrpAspIleSerGluSer-AlaLysAspPheIleCysHis-LeuLeuGln 259  
|||||  
Db 827 CTCGGTCTGGGATGACATTTCTGTAGTCAGGCCAGGGATTTTATTTGGCATCTCTCTGGA 886  
QY 259 uLysAspPro 262

Db 887 GAAGGACCCC 896  
|||||

RESULT 3  
BI821474  
LOCUS  
DEFINITION  
603038366F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5179336 5',  
mRNA sequence.  
ACCESSION  
BI821474  
VERSION  
BI821474.1 GI:159333024  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 809)  
AUTHORS  
NIH-MGC http://mgc.nci.nih.gov/  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished (1999)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM1447 row: c column: 17  
High quality sequence stop: 808.

FEATURES  
Location/Qualifiers  
1..809  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:5179336"  
/clone\_lib="NIH\_MGC\_115"  
/lab\_host="DH10B"  
/note="Organ: pooled brain, lung, testis; Vector:  
pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
source anonymous pool of 6 male brains, age range 23-27; 1  
male lung, age 27; and 1 male testis, age 69. Library is  
oligo-dT primed and directionally cloned (EcoRV site is  
destroyed upon cloning). Average insert size 1.8 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by C.  
Gruber (Invitrogen). Research Genetics tracking code  
021. Note: this is a NIH\_MGC Library."

BASE COUNT 212 a 192 c 218 g 187 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3,24e-118 Length: 809  
Score: 1257.00 Matches: 241  
Percent Similarity: 99.59% Conservative: 1  
Best Local Similarity: 99.18% Mismatches: 1  
Query Match: 50.02% Indels: 0  
DB: 13 Gaps: 0

US-09-960-643-2 (1-476) x BI821474 (1-809)

QY 1 MetGlyArgLysGluGluAspCysSerTrpLysLysGlnThrAsnIleArg 20  
|||||  
Db 75 ATGGGTGCAAGGAAGAGATGATGCTGCTGGAAGAAACACACCAACATCCCG 134  
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
|||||  
Db 135 AAAACCTTCATTTTATGGAAGTGTGGGATCAGGAGCTTCTCAGAAGTTTCTCTGGTG 194  
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60  
|||||  
Db 195 AAGCAAAAGACTGACTGGGAGGCTTGTGCTCTGAAAGTGCATCAAGAAGTCAACGCTGCTTC 254  
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80

Db	255	CGGGACAGCACCCTGGAGAATGAGATTGCTGTGGTGA AAAAGATCAACAGCATGA AACAATT	
Qy	81	ValThrLeuGluAspIleTy rGluSerThrThrHisTy rTy rLeuValMetGlnLeuVal	
Db	315	GTGACCCCTGGAGGCATCTATGAGACACACCCACTACTACCTGGTCATGCAGCTGTT	
Qy	101	SerGlyGlyGluLeuPheA spArgIleLeuGluAr gGlyValTy rTh rGluLysAspAla	
Db	375	TCTGTGGGGAGSCCTTTTGACCGGATCCTGGAGCGGGGTGCTACACAGAGAAGGATGCC	
Qy	121	SerLeuValIleGlnGlnValLeuSerAlaValLysTy rLeuHisGluAsnGlyIleVal	
Db	435	AGTCTGGTGATCCAGCAGGCTTTGTCGCAGTGA AATACTACATGAGATGGCATCGTC	
Qy	141	HisArgaspLeuLysProGluA snLeuLeuTy rLeuTh rProGluGluA snSerLysIle	
Db	495	CACAGAGACTTAAGACCCGAAAACCTGCTTTACCTTACCCCTGAAGAGAACTCTAAGATC	
Qy	161	MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys	
Db	555	ATGATCACTGACTTTGGTCTGCTCCAAGATGGACAGATGGCATCATGTCCACTGCCGTG	
Qy	181	GlyThrProGlyTy rValAlaProGluValLeuAlaGlnLysProTy rSerLysAlaVal	
Db	615	GGGACCCACAGGCTACGTGGCTCCAGAAGTCTGGGCCA AAACCCCTACAGCAAGGCTG	
Qy	201	AspCysTyrSerIleGlyValIleThrTy rIleLeuLeuCysGlyTy rProProPheTyr	
Db	675	GATTGCTGGTCCATCGCGGTCA TCACTACATAT TGCTGTGGATACCCCCGGTTCAT	
Qy	221	GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTy rTy rGluPheGluSer	
Db	735	GAAGAAGCGAGTCTAAAGCTTTTCGAGAA GATCCAGGAGGGGCTATTATGATTTGGAGTC	
Qy	241	ProPheTyrP 243	
Db	795	CCATTCTGG 803	
RESULT 4			
BI772626			
LOCUS			
DEFINITION		BI772626 812 bp mRNA linear EST 25-SEP 603060879f1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5210299	
ACCESSION		BI772626	
VERSION		BI772626.1 GI:15764204	
KEYWORDS		EST.	
SOURCE		human.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE		NIH-MGC http://mgc.nci.nih.gov/ 1 (bases 1 to 812)	
JOURNAL		National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)	
COMMENT		Contact: Robert Strausberg, Ph.D. Email: cgabs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LIAm11527 row: m column: 20 High quality sequence start: 2 High quality sequence stop: 808.	

FEATURES  
SOURCE

/lab\_host="DH10B"  
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;  
Site\_1: NotR; Site\_2: EcoRV (destroyed); RNA source:  
anonymous pool of 24 week female lung, 16 week female  
spleen, and 20-22 week male spleens. Library is oligo-dT  
primed and directionally cloned (EcoRV site is destroyed  
upon cloning). Average insert size 1.4 kb, insert size  
range 1-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 026. Note:  
this is a NIH MGC library."

	217 a	186 c	221 g	188 t
BASE COUNT				
ORIGIN				

Alignment Scores:

Alignment Scores:	6.62e-116	Length:	812
Pred. No.:	1234.50	Matches:	242
Score:	1234.50	Conservative:	1
Percent Similarity:	90.00%	Mismatches:	2
Best Local Similarity:	89.53%	Indels:	26
Query Match:	49.12%	Gaps:	1
DB:	13		

US-09-960-643-2 (1-476) x BI772626 (1-812)

Qy	1	MetGlyArgLysGluGluAspAspCysSerSerTrpLysGlnThrThrAsnIleArg	20
Db	78	ATGGGTCAAGAGGAAAGAGATGACTGCAGTTCTCTGGAGAAACAGACCACCAACATPCCGG	137
Qy	21	LysThrPheIlePheMetGluValLeuGlySerGlyValaPheSerGluValPheLeuVal	40
Db	138	AAACCTTCATTTTATGGAAGTCTGGGATCAGGACITTCACAGAAGTTTCTCGGTG	197
Qy	41	LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe	60
Db	198	AAGCAAAAGACTGACTGGGAAGCTTTTGCTCTGAAGTGCATCAAGAAGTCACCTGCCTTC	257
Qy	61	ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle	80
Db	258	CGGACAGACACCTGGAGAAATGAGATGCTGTCTGAAANAAGATCAAGCATGNAACATT	317
Qy	81	ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal	100
Db	318	GTGACCTGGAGGACATCTATGAGAGCACCACCACCTACTACCTGTGTCATGCAGCTTGT	377
Qy	101	SerClyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla	120
Db	378	TCGTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTCTTACACAGAGAGATGCC	437
Qy	121	SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal	140
Db	438	AGTCTGTGTATCCAGCAGGCTTGTTCGGCAGTCAAAATACCTACATCAGAAATGGCATCGTC	497
Qy	141	HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluAsnSerLysIle	160
Db	498	CACAGAGACTTAAGCCCCGAAACCTTCCTTACCTTACCCCTGAAGAGAACTCTAGATC	557
Qy	161	MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys	180
Db	558	ATGATCACTGACTTGGCTCTGTCCAAGATGGAACAGAAATGGCATCATGTCCACATGCCTGT	617
Qy	181	GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal	200
Db	618	GGGACCCCGAGCTACGGGG-----	637
Qy	201	AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProProPheTyr	220
Db	638	-----CTCTGTGGATACCCCCATCTCTAT	661
Qy	221	GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrGluPheGluSer	240
Db	662	GAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTACTATGAGTATGAGTCT	721
Qy	241	ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys	260

```
Db 722 CCATTCTGGGATGACATTTCTGAGTCAGCCAGGAGCTTTATTGGCCACTTGTCTTGAGAAG 781
QY 261 AspProAsnGluArgTyrThrCysGluLys 270
Db 782 GATCCGACGAGCGGTACACCTGTGAGAAG 811

RESULT 5
BM807335
LOCUS BM807335 1129 bp mRNA linear EST 05-MAR-2002
DEFINITION AGENCOURT 6575096 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5732298
5' UTR sequence.
ACCESSION BM807335
VERSION BM807335.1 GI:19124158
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (Bases 1 to 1129)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgs@ncl.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM2734 row: c column: 19
High quality sequence start: 11
High quality sequence stop: 651.
Location/Qualifiers
1. 1129
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5732298"
/clone_lib="NIH_MGC_124"
/tissue_type="hippocampus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site: 1: EcoRV
(destroyed); Site: 2: NotI; RNA source male hippocampus,
age 27; library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."
BASE COUNT 287 a 308 c 294 g 240 t
ORIGIN

Alignment Scores:
Pred. No.: 2.89e-111 Length: 1129
Score: 1191.50 Matches: 281
Percent Similarity: 72.86% Conservative: 9
Best Local Similarity: 70.60% Mismatches: 40
Query Match: 47.41% Indels: 70
DB: 14 Gaps: 7

US-09-960-643-2 (1-476) x BM807335 (1-1129)
QY 1 MetGlyArgLysGluGluAspAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
Db 105 ATGGGTGCAAGGAGGAGATGACTGCAGTTCCTGGGAAGAACAGACCAACATCCCG 164
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 165 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAAGTTTCTCGGTG 224
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60

Db 225 AAGCAAAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCAACCTGCCTTC 284
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysLysLysHisGluAsnIle 80
Db 285 CGGACAGACAGCCTGGAGAAATGAGATTGCTGTGTTGAAAAGATCAAGCATGAAACATT 344
QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
Db 345 GTGACCTGGAGGACATCTATGAGAGCACCACCACTACTACTGCTGGTCAATGCACTGTGT 404
QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
Db 405 TCTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGATGCC 464
QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
Db 465 AGTCTGGTATCCAGCAGGCTTGTTCGGCAGTGAATACCTACATGAGATGCGATCGTC 524
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrThrProGluGluAsnSerLysIle 160
Db 525 CACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCTGAAGAGAACTCTAAGATC 584
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
Db 585 ATGATCACTGACTTTGGTCTGTCCAAGATGAACAGATGCGATCATGCTCCACTGCTGT 644
QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAla-Va 200
Db 645 GGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCCGAGAAACCTTACAGCAAGGCTGT 704
QY 200 Lasp-CysTrp-SerIleGlyValIleThr-TyrIleLeuLeuCysGly---TyrProPr 218
Db 705 GGATTGCTGGTCCATCGGGGGCATCACCTACCTATTGCTCTGGTAAACCCGCCCC 764
QY 218 oPheTyrGlu-GluThrGluSerLysLeuPheGluLysIleLysGluGlyTyr---TyrG 237
Db 765 GTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAAGGGCTAATATGA 824
QY 237 LuPheGluSerProPheTrp-AspAsp-IleSerGluSerAla-LysAspPhe-IleCys 255
Db 825 GTTTTGATCTCCATTCTGGGGATGAACATTCCTGAGTCAGCCCGAGGAGCTTTAATTTGC 884
QY 256 HisLeu---LeuGluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHis 274
Db 885 CACCTTGGCTTGAAGAGGATCCCAACCGAAGGGGGACCCCGG----- 929
QY 275 ProTrpIleAspGlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGln 294
Db 930 -----GGGAACAGGCGCTTTGAG----- 947
QY 295 IleGlnLysAsnPheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValVal 314
Db 947 ----- 947
QY 315 HisHisMetArgLysLeuHisMetAsnLeuHisSerProGlyValArgProGluValGlu 334
Db 948 -----CACTCCCGGGGACA-TGGACGGGAAACACT 976
QY 335 AsnArgProPro-----GluThrGlnAlaSerGluThrSerArgProSerProGlu 352
Db 977 GGCCCTCCCGGGGGGCCACTCCCGCTCTCATATAAGCCT-CCCCAAAATCCAGAG 1035
QY 353 IleThrIleThrGluAlaProValLeuAspHisSerValAlaLeuPro-AlaLeuThrG1 372
Db 1036 AAAACTTTGTCC-----ATACCACACCAAGTGGCGACGCCCAAGCGCTTTAAC 1083
QY 372 nLeuProCysGlnHisGlyArgArgProThrAla 383
Db 1084 GGCCCGCCAGGACTCGTGGCGCGGCCCGCCACTTCT 1117
RESULT 6
AL560091
```

LOCUS AL560091 1051 bp mRNA linear EST 16-FEB-2001  
 DEFINITION LTI\_FL011\_Bc1 Homo sapiens cDNA clone CS0DG002YN16 5 prime  
 , mRNA sequence.

ACCESSION AL560091  
 VERSION AL560091.1 GI:12906218  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1051)

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. 1051

source

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DG002YN16"

/clone\_lib="LTI\_FL011\_Bc1"

/sex="male"

/tissue\_type="B cells from Burkitt lymphoma"

/lab\_host="DH10B"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-stranded cDNA was digested with Not I and cloned  
 into the Not I and Eco RV sites of the pCMVSPORT 6 vector.

Library was constructed by Life Technologies. Contact :

Feng Liang Life Technologies, a division of Invitrogen

9800 Medical Center Drive Rockville, Maryland 20850, USA

Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 274 a 270 g 205 t 15 others

ORIGIN

Alignment Scores:

Pred. No.: 7.02e-108 Length: 1051

Score: 1158.00 Matches: 226

Percent Similarity: 84.29% Conservative: 37

Best Local Similarity: 72.44% Mismatches: 47

Query Match: 46.08% Indels: 3

DB: 9 Gaps: 2

US-09-960-643-2 (1-476) x AL560091 (1-1051)

Qy 1 MetGlyArgLysGluGluAspCysSerTrpLysLysGlnThrThrAsnIleArg 20

Db 111 ATGCCCGGGAGACGGCGAGAGACGCTCTCTGGAAAAGCAAGCTGAAGACATCAAG 170

Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40

Db 171 AAGATCTTCGAGTTCAAGAGACCTCGAACCAGGGCCCTTCCGAAGTGGTTTAGCT 230

Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59

Db 231 GAAGAGAGGCAACTGGCAAGCTTTTCTGTGAAGTGTATCCCTAAGAGGGCGCTGAAG 290

Qy 60 PheArgAspSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHIsGluAsn 79

Db 291 GCGAAGGAAACACATAGAGATGAGATAGCGCTCTGAGAAGATTAAGCATGAAGAT 350

Qy 80 IleValThrLeuGluAspIleTyrgLuserThrThrHisTyrtYrLeuValMetGlnLeu 99

Db 351 ATTTGTCCTCGGAAGACATTTATGAAAGCCCAATCACCTGATCTGGTCATGCAGCTG 410

Qy 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyThrGluLysAsp 119

Db 411 GTGTCGGTGGAGAGCTGTTTGACCGGATAGTGGAGAGGGGTTTATACAGAGAGGAT 470

Qy 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrrLeuHisGluAsnGlyIle 139

Db 471 GCCAGCATCTGATCCGCCAAGTCTTGACGCGGTGTACTATCTCCACAGAATGGCATC 530

Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrrLeuThrProGluGluAsnSerLys 159

Db 531 GTCCACAGAGACCTCAAGCCGAAATCTCTGTACTACAGTCAAGATGAGAGATCCAAA 590

Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178

Db 591 ATAATGATCAGTACTTTGGATGTCAAAATATGGAGGCAAGAGATGTGATGTCCACT 650

Qy 179 AlaCysGlyThrProGlyTyrrValAlaProGluValLeuAlaGlnLysProTyrrSerLys 198

Db 651 GCCTGTGGAACCTCCAGGCTATGTGCTCTCTGAAGTCCTCGCCACAGAAACCTTACAGCAA 710

Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrrIleLeuLeuCysGlyTyrrProPro 218

Db 711 GCGGTGTACTGCTGTGTCATCGGAGTGATGCTTGTCTTGGCTGCGGCTACCCGCT 770

Qy 219 PheTyrrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrrTyrrGluPhe 238

Db 771 TTTTATGATGAATGACTCCAAGCTCTTGAGCAGATCCTCAAGCGCGAATATGAGTTT 830

Qy 239 GluSerProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258

Db 831 GACTCTCCCTACTGGGTGACATCTCCGACTCTGCAAAAGACTTTCATTCG-GAATGATG 889

Qy 259 GluLysAspProAsnGluArgTyrrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278

Db 890 GAGAAGGCCCGAATAAAGATACACGTGTGAGCAGGAGCTCGGCCACCCATGATCGCT 949

Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrrProSerValSerLeuGlnIleGlnLysAsn 298

Db 950 GTGACACAGCCCTCAAAATAAACATCCAGGAGTCCGTCAGCGCCAGATCGGAAAAAC 1009

Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAla 310

Db 1010 TTTGCCAAGAGCAATGGAGACAGCATTTTATGCC 1045

RESULT 7

BI824483 BI824483 740 bp mRNA linear EST 04-OCT-2001

LOCUS 603038855F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5179957 5',

DEFINITION mRNA sequence.

ACCESSION BI824483

VERSION BI824483.1 GI:15936033

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 740)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Contact: cgabs@remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM11448 row: m column: 14

High quality sequence stop: 709.

Location/Qualifiers

1. 740

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5179957"

/clone\_lib="NIH\_MGC\_115"

FEATURES

SOURCE

```
/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
pCMW-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT      196 a   181 c   196 g   167 t
ORIGIN

Alignment Scores:
Pred. No.:      5,37e-107      Length:      740
Score:          1147.00      Matches:      222
Percent Similarity: 99.55%      Conservative: 0
Best Local Similarity: 99.55%      Mismatches: 1
Query Match:      45.64%      Indels:      0
DB:              13      Gaps:      0

US-09-960-643-2 (1-476) x BI824483 (1-740)

QY 1 MetGlyArgLysGluGluAspAspCysSerTrpLysLysGlnThrThrAsnIleArg 20
Db 70 ATGGGTCGAAGGAAGAAGATGACTGCTGCTGGAAGAAACACACACCACCAATCCGG 129
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 130 AAAACCTTCATTTTATGGAAGTGTGGGATCAGGAGCTTCTCAGAAAGTTTCTCGTGTG 189
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
Db 190 AAGCAAGACTGACTGGGAGCTTTGTCTCTGAAGTGCATCAAGAAAGTCACCTCCCTTC 249
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysLysIleLysHisGluAsnIle 80
Db 250 CGGACAGACGCTGGAGATGAGATGCTGTGTGTTGAAAAAGATCAAGCATGAAACATT 309
QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
Db 310 GTGACCTGTGGAGACATCTATGAGAGCACCCACCTACTACTGCTGTCATGCGAGTTGTT 369
QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
Db 370 TCTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTCTACACAGAGAGGATGCC 429
QY 121 SerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
Db 430 AGTCTGTGATCCAGCAGGCTTTGTGCGCAGTGAATACTACATGAGAATGGCATCGTC 489
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
Db 490 CACAGAGACTTAAGCCCGAAAAACCTGCTTACCTTACCCCTGAAGAGAACTCTAAGATC 549
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
Db 550 ATGATCACTGACTTTGTGCTGTCCAAGATGGAACAGATGGCATCATGTCCTCCACTG 609
QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Db 610 GGGACCCAGGCTACGTGGCTCCAGAAGTGCTGGCCCGAAAAACCCCTACACAGAGGCTGTG 669
QY 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLysCysGlyTyrProProPheTyr 220
Db 670 GATTGCTGTTCATCGGCGTCATCACTACATATTGCTCTGGGATACGCCCCCGTTCTAT 729
QY 221 GluGluThr 223
Db 730 GAAGAAACG 738

RESULT 8
```

```
BM921532
LOCUS
DEFINITION
AGENCOURT_6708041 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5753010
5', mRNA sequence.
ACCESSION
BM921532
VERSION
BM921532.1 GI:19371911
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1068)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cyapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12788 row: b column: 19
High quality sequence stop: 592.
Location/Qualifiers
location=1..1068
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5753010"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
pCMW-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT      279 a   281 c   260 g   248 t
ORIGIN

Alignment Scores:
Pred. No.:      4.49e-106      Length:      1068
Score:          1140.50      Matches:      229
Percent Similarity: 94.67%      Conservative: 2
Best Local Similarity: 93.85%      Mismatches: 10
Query Match:      45.38%      Indels:      3
DB:              14      Gaps:      1

US-09-960-643-2 (1-476) x BM921532 (1-1068)

QY 1 MetGlyArgLysGluGluAspAspCysSerTrpLysLysGlnThrThrAsnIleArg 20
Db 83 ATGGTCAAGGAAGAAGATGACTGCTGGAAGAAACACAGACCACCAATCCGG 142
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 143 AAAACCTTCATTTTATGGAAGTGTGGGATCAGGAGCTTCTCAGAAGTTTTCCTGGTG 202
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
Db 203 AAGCAAGACTGACTGGGAGCTTTCCTCTGAAGTGCATCAAGAGATCACCCTGCTTC 262
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
Db 263 CGGACAGCAGCCCTGGAGATGAGATTGCTGTGTGAAAAAGATCAAGCATGAACAT 322
QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
```

```

|||||
Db 323 GTGACCTCGAGGACATCTATGAGAGCACCACCTACTACCTGGTCATGCGCTTGT 382
QY 101 SerGlyGlyLeuPheAspArgLeuGluArgGlyValThrGluLysAspAla 120
Db 383 TCTGTGGGGAGCTCTTTACCGGATCCCTGGAGCGGGTGTCTACACAGAACGATGCC 442
QY 121 SerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
Db 443 AGTCTGGTATCCAGCAGGTCTGTGGCAGTGAATACCTACATGAGAATGGCATCGTC 502
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
Db 503 CACAGAGACTTAAAGCCGAAAACTGCTTTACCTTACCCCTGAAGAGAACTCTTAAGATC 562
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGluAsnGlyIleMetSerThr-AlaCys 180
Db 563 ATGATCACTGACTTTGGTGTCTGCAAGATGGAACAGATGGCATGATGCCACCTGCCTG 622
QY 180 sGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Db 623 TGGACCCCGAGCTAGCTGGCTCCAGAGTGTGGCCAGAAACCTACAGCAAGGCTGT 682
QY 200 lAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPhe-T 220
Db 683 GGATTGCTGCTCCATCGGGGTATCACCCTACATATTGCTCTGGGATACCCCGCTCCT 742
QY 220 yrGluGluThrGluSerLysLeuPhe--GluLysIleLysGluGlyTyrGluPheG 239
Db 743 ATGAAGAAGCGAATCTAACCTTTTCGAGAAGAATCAAGGAGGCGCTACTATGAGATTC 802
QY 239 luSerPro 241
Db 803 TGAGCCCC 810

RESULT 9
BI1818261
LOCUS 603032510F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173587 5',
DEFINITION mRNA sequence.
ACCESSION BI1818261
VERSION BI1818261.1 GI:15928724
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 742)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1432 row: d column: 04
High quality sequence stop: 742.
Location/Qualifiers
1..742
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5173587"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is

```

## FEATURES

source

1..742

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5173587"

/clone\_lib="NIH\_MGC\_115"

/lab\_host="DH10B"

/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is

oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH\_MGC Library."

BASE COUNT 197 a 180 c 197 g 168 t

## ORIGIN

Alignment Scores: 1.16e-105 Length: 742  
Pred. No.: 1134.00 Matches: 220  
Score: 1134.00 Conservative: 0  
Percent Similarity: 99.10% Mismatches: 2  
Best Local Similarity: 99.10% Indels: 0  
Query Match: 45.13% Gaps: 0  
DB: 13

US-09-960-643-2 (1-476) x BI1818261 (1-742)

QY 1 MetGlyArgLysGluGluAspAspCysSerSerTriplyLysGlnThrThrAsnIleArg 20  
Db 77 ATGGTCTGAAGGAAGAAGATGACTGCTGAGTCTCTGGGAAGAAACAGACCACCAACATCCGG 136  
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
Db 137 AAAACCTTCAATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTCTTCTCGGTG 196  
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60  
Db 197 AAGCAAGACTGACTGGGAAGCTCTTTCCTGCTGAGTGCATCAAGAAGTCACTCGCTTC 256  
QY 61 ArgAspSerLeuGluAsnGluIleAlaValLeuLysLysLysHisGluAsnIle 80  
Db 257 CGGCACAGCAGCCTGGAGAAATGAGATTGCTGTGTGAAAAAGATCAAGCATGAAACATT 316  
QY 81 ValThrLeuGluAspIleTyrGluSerThrHisTyrTyrLeuValMetGlnLeuVal 100  
Db 317 GTGACCTGGAGGACATCTATGAGAGCACCACCACCTACTACCTGGTGCATGCGCTTGT 376  
QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120  
Db 377 TCTGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGGATGCC 436  
QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140  
Db 437 ACTCTGTGATCCAGCAGGCTTCTGCGCAGTGAATACCTACATGAGAATGGCATCGTC 496  
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160  
Db 497 CACAGAGACTTAAAGCCGAAAACTGCTTTACCTTACCCCTGAAGAGAACTCTAAGATC 556  
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180  
Db 557 ATGATCACTGACTTTGGTCTCTCAAGATGGAACAGATGGCATCATGTCCATCGCTGT 616  
QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200  
Db 617 GGGACCCCGAGCTACGTGGCTCCAGAAAGTGTGGCCAGAAAGCTACAGAGGCTGTG 676  
QY 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyr 220  
Db 677 GATTGCTGGTCCATCGGCGCTCATCACCTACATATAGTCTCTGGATACCCCGCTTCTAT 736  
QY 221 GluGlu 222  
Db 737 GAAGAA 742

## RESULT 10

BI753035

## LOCUS

603025844F1 NIH\_MGC\_114

## DEFINITION

mRNA sequence.

## ACCESSION

BI753035

BI753035 731 bp mRNA linear EST 25-SEP-2001  
603025844F1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5196323 5',  
mRNA sequence.

```
VERSION      B1753035.1  GI:15744613
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.
TITLE        1 (bases 1 to 731)
JOURNAL      NIH-MGC http://mgc.nci.nih.gov/.
COMMENT      National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: csaapbs-r@mail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LUAM11491 row: g column: 12
              High quality sequence stop: 729.
FEATURES     1. .731
              Location/Qualifiers
              source
                1. .731
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5196323"
                /lab_host="DH10B"
                /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                male brains, age range 23-27 yo. Library is oligo-dT
                primed and directionally cloned (EcoRV site is destroyed
                upon cloning). Average insert size 1.5 kb, insert size
                range 1-3 kb. Library is normalized and enriched for
                full-length clones and was constructed by C. Gruber
                (Invitrogen). Research Genetics tracking code 019. Note:
                this is a NIH_MGC Library."
BASE COUNT   190 a 182 c 191 g 168 t
ORIGIN
Alignment Scores:
Pred. No.:    1.93e-104    Length:    731
Score:        1122.00     Matches:    217
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match:   44.65%      Indels:    0
DB:           13          Gaps:    0
US-09-960-643-2 (1-476) x B1753035 (1-731)
QY    2  GlyArgLysGluGluAspAspCysSerSerTriplyLysGlnThrThrAsnIleArgLys 21
DB    73  GGTGGAAGGAGGAGATGACTGCTGCTGGGAGAACAGACACCAACATCCGGAAA 132
QY    22  ThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuValLys 41
DB    133  ACCCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAAGTTTCTCTGGTGAAG 192
QY    42  GlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPheArg 61
DB    193  CAAGACGTGACTGGGAGCTCTTGGCTCTGAAGTGCATCAAGAGTCACTGCTGCTCCCG 252
QY    62  AspSerSerLeuGluAsnGluIleAlaValLeuLysIleLysHisGluAsnIleVal 81
DB    253  GACAGCAGCCCTGGAGATGAGATGCTGTGTGAAAAAGATCAAGCATGAAACATTTGTG 312
QY    82  ThrLeuGluAspIleTyrGluSerThrHisTyrTyrLeuValMetGlnLeuValSer 101
DB    313  ACCCTGGAGGACATCTATGAGAGACCAACCACTACTACCTGCTGCTGCTGCTGCTGCT 372
QY    102  GlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAlaSer 121
DB    373  GGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTCTTACAGAGAGGATGCCAGT 432
```

```
QY    122  LeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleValHis 141
DB    433  CTGGTGATCCAGCAGGCTCTTCGGCACTGAAATACCTACATGAGAAATGGCATCGTCCAC 492
QY    142  ArgAspLeuLysProGluAsnLeuLeuTyrThrProGluGluAsnSerLysIleMet 161
DB    493  AGAGACTTTAAAGCCGGAACCTGCTTACCTTACCCCTGAAGAGAACTCTTAAGATCATG 552
QY    162  IleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCysGly 181
DB    553  ATCACTGACTTGGTCTGTCCAGATGGAACAGATGSCATCATGTCCTGCTGCTGGG 612
QY    182  ThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaValAsp 201
DB    613  ACCCCAGGCTACGTGGCTCCAGAAAGTGTGGCCAGAAACCTACAGCAAGGCTGTGGAT 672
QY    202  CysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
DB    673  TGCTGTCATCGCGCTCATCACCTACATATTTGCTGTGTGGATACCCCCCC 723
RESULT 11
BG293660      818 bp      mRNA      linear      EST 21-FEB-2001
LOCUS        602390529F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4502479 5',
DEFINITION   mRNA sequence.
ACCESSION    BG293660
VERSION      BG293660.1  GI:13053536
KEYWORDS     EST.
SOURCE       house mouse.
ORGANISM     Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheraia; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 818)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: csaapbs-r@mail.nih.gov
              Tissue Procurement: The Cepko Laboratory
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LUAM10371 row: i column: 08
              High quality sequence stop: 765.
FEATURES     1. .818
              Location/Qualifiers
              source
                1. .818
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /clone="IMAGE:4502479"
                /clone_lib="NIH_MGC_94"
                /tissue_type="retina"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                Average insert size 3.3 kb. Library enriched for
                full-length clones and constructed by Life Technologies.
                Note: this is a NIH_MGC Library."
BASE COUNT   216 a 191 c 221 g 190 t
ORIGIN
Alignment Scores:
Pred. No.:    2.05e-102    Length:    818
Score:        1103.00     Matches:    228
Percent Similarity: 94.69%    Conservative: 4
Best Local Similarity: 93.06%    Mismatches: 11
Query Match:   43.89%      Indels:    5
DB:           12          Gaps:    0
US-09-960-643-2 (1-476) x BG293660 (1-818)
```

```
QY 1 MetGlyArgLysGluGluAspLysCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
|||||
Db 86 ATGGGCGTAAGGAGGAGGAGGAGTGCAGTCTCTGGAGAAACACAGACCACCAATCAGG 145
|||||
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
|||||
Db 146 AAAACCTTCATCTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAGGTCTCTGGTG 205
|||||
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
|||||
Db 206 AAGCAAGAGTGACTGGAAACTCTTTGCTCTGAAATGATATCAAGAAGTCACCAAGCTTC 265
|||||
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
|||||
Db 266 CGGGACAGCAGCCTAGAGATGAGATCGCTGTCTGAAAGGATCAAGCATGAGAATCATT 325
|||||
QY 81 ValThrLeuGluAspIleTrpGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
|||||
Db 326 GTGACCTCGGAGGACATCTATGAGACACACCCACTACTACTGCTGTCATGAGCTTGT 385
|||||
QY 101 SerGlyGlyLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
|||||
Db 386 TCTGGAGGTGAGCTTTTACCGGATCTTAGAGCGGTGTCTACACAGAAAAGGATGCC 445
|||||
QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
|||||
Db 446 AGCCTGTCTATCCAGCAGGTCTGTCTGCGGTGAAATCTTCATGAGAAATGGCATCGTC 505
|||||
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
|||||
Db 506 CACAGAGACTAAAGCCTGAAACCTGCTGTACCTACCCCTGAGGAGAACTCCAGATC 565
|||||
QY 161 MetIleThrAspPhe-GlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCy 180
|||||
Db 566 ATGATCAGTCTGTGTGTCTATCCAAAGATGAGCAGAAATGGAGTCATGTC-ACAGCTTG 624
|||||
QY 180 sGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVa 200
|||||
Db 625 TGGAGCCCCAGGTACGTGGCTCCAGAGTGTGGCCCCAGAGCCCTCAGTAAGGCTGT 684
|||||
QY 200 LysCysTrpSerIleGlyValIleThrTyrIleLeuLeuLysGlyTyrProPro- 220
|||||
Db 685 GGACTGCTGTC-ATTGGTGTATCATCATACATACATGTC-TGTGGGTATCCCTTTTCT 742
|||||
QY 220 YrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluS 240
|||||
Db 743 ATGAAGAACAGAATCAAAGTTTTTCGAGACCATCATCAGAGGGTTACTACGAAGTTTGAA 802
|||||
QY 240 erProPheTrp 243
|||
Db 803 GTTCCGCTCTGG 813
|||||
```

```
RESULT 12
BG715920 824 bp mRNA linear EST 08-MAY-2001
LOCUS 602676667F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4799224 5',
DEFINITION mRNA sequence.
ACCESSION BG715920
VERSION BG715920.1 GI:13995107
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 824)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
```

```
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10688 row: e column: 17
High quality sequence stop: 766.
FEATURES
    source
        1..824
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4799224"
            /clone_lib="NIH_MGC_96"
            /tissue_type="hypothalamus"
            /lab_host="DH10B"
            /note="Organ: brain; Vector: pBluescriptR (modified
            pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
            ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
            size-selected for average insert size 2.3 kb and
            normalized to ROT 5. This is a primary library enriched
            for full-length clones and constructed using the
            Cap-trapper method (Carninci, in preparation). Library
            constructed by M. Brownstein (NIH/NHGRI, National
            Institutes of Health). Note: this is a NIH_MGC Library."
```

```
BASE COUNT 223 a 194 c 219 g 188 t
ORIGIN
Alignment Scores:
Pred. No.: 3.33e-102 Length: 824
Score: 1101.00 Matches: 222
Percent Similarity: 96.58% Conservative: 4
Best Local Similarity: 94.87% Mismatches: 6
Query Match: 43.81% Indels: 3
DB: 12
US-09-960-643-2 (1-476) x BG715920 (1-824)
```

```
QY 1 MetGlyArgLysGluGluAspLysCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
|||||
Db 107 ATGGTCTCAAAGGAGGAGGAGTGCAGTCTCTGGAGAAACACAGACCACCAATCGG 166
|||||
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
|||||
Db 167 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGATTTCTCGTG 226
|||||
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
|||||
Db 227 AAGCAAGAGCTGACTGGGAAGCTCTTCTGCTGAAGTGCATCAAGAAGTCACCTGCCTTC 286
|||||
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
|||||
Db 287 CGGGACAGCAGCTGGAGAAATGAGGTCTGTGTTGAAAGAGATCAAGCATGAAACATT 346
|||||
QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
|||||
Db 347 GTGACCTCGGAGGACATCTATGAGACACCAAGCTACTACTCTGTCATGAGCTTGT 406
|||||
QY 101 SerGlyGlyLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
|||||
Db 407 TCTGTGGGGAGCTCTTTTACCGGATCATGAGCGGGGTCTACACAGAGAAGATGCC 466
|||||
QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
|||||
Db 467 AGTCTGGGTATCCAGCAGGTCTGTGCGCAGTGAATACCTACATGAGAATGGCATCGTC 526
|||||
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
|||||
Db 527 CACAGAGACTTAAAGCCGAGAAACCTGCTTTACCTTACCTCCCTGAAGAGAACTCTAAGATC 586
|||||
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
|||||
Db 587 ATGATCAGTACTGTTGGTCTGTCCAAAGATGGAACAGAAATGGCATCATGCTCCACTGCTGT 646
|||||
```

Qy 181 GlyThrProGlyTyrVal-AlaProGluValLeuAlaGln-LysProTyrSerLysAlaV 200  
 Db 647 GGGACCCAGGCTACGTGGCTGCAGAGTGTGGCCGAGAAACCTACAGCAGCG 706  
 Qy 200 aAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheT 220  
 Db 707 TGGATTCTGGTCCATCGGGGT-ATCACCTAACATATGCTCTGTGGATACCCCCATCT 765  
 Qy 220 yrGluGluThrGluSerLysLeuPheLysLysLys 232  
 Db 766 ATGAAGAACGAGCTCTTACGCTTTTCGAAAAATCAAG 803

RESULT 13  
 LOCUS BQ949629 918 bp mRNA linear EST 21-AUG-2002  
 DEFINITION AGENCOURT\_8744314 Lupski\_sciatic\_nerve Homo sapiens cDNA clone  
 IMAGE:6205867 5', mRNA sequence.  
 ACCESSION BQ949629  
 VERSION BQ949629.1 GI:22365107  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 918)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: Dr. James R. Lupski  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM13629 row: g column: 20  
 High quality sequence stop: 697.

FEATURES  
 source  
 1..918  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6205867"  
 /lab\_host="Lupski\_sciatic\_nerve"  
 /sex="male"  
 /tissue\_type="sciatic nerve"  
 /dev\_stage="adult, 70 yr"  
 /lab\_host="PH10B"  
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site\_1:  
 NotI; Site\_2: SalI; cDNA made by oligo-dT priming.  
 Directionally cloned using the following adaptors:  
 5'-TCGACCCAGCGCTCG-3' and  
 5'-GACTATTCTAGATCCGAGCGCGCCCTT(15)-3'. Size selected >  
 1 kb for average insert length 1.87 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine) and is available through Life  
 Technologies."

BASE COUNT 237 a 240 c 250 g 191 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3,73e-93 Length: 918  
 Score: 1013.50 Matches: 200  
 Percent Similarity: 82.86% Conservative: 32  
 Best Local Similarity: 71.43% Mismatches: 43  
 Query Match: 14.33% Indels: 5  
 DB: 14 Gaps: 3

US-09-960-643-2 (1-476) x BQ949629 (1-918)

Qy 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20  
 Db 79 ATGGCCCGGAGAACGGGAGAGAGAGCTCCCTCGGAAAAAGACAGCTGAAGACATCAAG 138  
 Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
 Db 139 AAGATCTTCGAGTTCAAAGACACCTCGGAACCGGGGCTTTCCGAAAGTGGTTTAACT 198  
 Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAla 59  
 Db 199 GAAGAGAAGCAACTGGCAAGCTCTTGTCTGAAGTGTATCCCTAAGAAGCGCTGAAG 258  
 Qy 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79  
 Db 259 GGCAGGAAGAACGACATAGAAATGAGATACCCCTCGTGAAGAAGATTAAAGCATGAAGAT 318  
 Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99  
 Db 319 ATTGTGCTCCCTGGGAAGACATTTATGAAGCCCAATCACCTGTACTTGGTCATGCAGCTG 378  
 Qy 100 ValSerGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119  
 Db 379 GTGTCCGCTGGAGAGCTTTTACCGGATAGTGGGAAGGGGTTTATACAGAGAGGAT 438  
 Qy 120 AlaSerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139  
 Db 439 GCCAGCACTCTGATCCGCCAAGTCTTGGACCCGCTGACTATCTCCACAGATGGGATC 498  
 Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159  
 Db 499 GTCCACAGAGACCTCAAGCCCGAAATCTCTTGTACTACAGTCAAGATGAGGAGTCCAAA 558  
 Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThr 178  
 Db 559 ATAATGATCACTGACTTTGGATTGTCAAAATGGAGGCAAGGAGATGTGATGTCACCT 618  
 Qy 179 AlacysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198  
 Db 619 GCCTGTGGAACCTCCAGGCTATGCTGCTGAAGTCTCGCCCAAGAAACCTTACAGCAA 678  
 Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuLysGlyTyrProPro 218  
 Db 679 GCCTGTGACTCTGCTGCTCCATCGAGTATTGCTCTACATCTTGTCTCGGCTACCTCTCT 738  
 Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysLysGluGlyTyrTyrGluPhe 238  
 Db 739 TTTTATGATGAATGACTCCAGCTCTTTGAGCAGATCCCTCAGCGCGGATATGAGTTT 798  
 Qy 239 GluSerProPheTrp-AspAspIleSerGluSerAlaLysAspPheIleCysHis-LeuL 258  
 Db 799 GACTCTCCCTACTGGGGATGACATCTCCGACTCTGGCAAGACTTCATTCGGAACCTGA 858  
 Qy 258 euGluLysAspProAsnGluArgTyrThrCysGluLysAla---LeuSerHisPro 275  
 Db 859 TGGAGAAGGCCGCAATAAAGATACCCGCTGTGAAGCAGCAGCAGCTCGGCACCC 914

RESULT 14  
 LOCUS AW251224 555 bp mRNA linear EST 17-DEC-1999  
 DEFINITION UI-R-BJ0-adf-f-07-0-UI.s1 UI-R-BJ0 Rattus norvegicus cDNA clone  
 AW251224  
 VERSION AW251224.1 GI:6594815  
 KEYWORDS EST.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (bases 1 to 555)  
 AUTHORS Ronaldo, M.F., Lennon, G. and Soares, M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery

JOURNAL  
MEDLINE  
COMMENT

Genome Res. 6 (9), 791-806 (1996)  
97044477  
Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu  
Oligo-dT track not found, Not 1 site shown in beginning of sequence  
is likely internal to the message. cDNA Library Preparation: M.B.  
Soares Lab Clone distribution: clones will be available through  
Research Genetics (www.resgen.com)  
Seq primer: M13 Forward  
POLYA-No.

FEATURES  
source

Location/Qualifiers  
1..555  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-BJ0-adv-f-07-0-UI"  
/clone\_lib="UI-R-BJ0"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not 1; Site\_2: Eco RI; The UI-R-BJ0  
library is a subtracted library derived from the UI-R-AAL,  
UI-R-ABI, UI-R-ACL, UI-R-AD1, UI-R-AEL, UI-R-AFL, and  
UI-R-AG1 libraries. These libraries represent tissues from  
rat atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal  
at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV  
canal at 15 dpc, and ventricle at 13 dpc. The tag is a  
string of 5-6 nucleotides present between the Not 1 site  
and the oligo-dT track. The library was constructed as  
described by Bonaldo, Lennon and Soares, Genome Research  
6: 791-806, 1996.  
TAG\_LIB=UI-R-BJ0  
TAG\_TISSUE=AV canal at 16.5 dpc  
TAG\_SEQ=GAACC"

BASE COUNT 131 a 134 c 139 g 151 t  
ORIGIN

## Alignment Scores:

Pred. No.: 1.4e-90 Length: 555  
Score: 985.00 Matches: 181  
Percent Similarity: 98.91% Conservative: 1  
Best Local Similarity: 98.37% Mismatches: 2  
Query Match: 39.20% Indels: 0  
DB: 10 Gaps: 0

US-09-960-643-2 (1-476) x AW251224 (1-555)

Qy 128 LeuSerAlaValLysTyrLeuHisGluAnGlyIleValHisArgAspLeuLysProGlu 147  
Db 553 TTGTCGCGGTGAAATACCTTCATGAGAAATGGCATCGTTCCACAGAGATCAAGGCTGAA 494  
Qy 148 AsnLeuLeuTyrLeuThrProGluGluAnSerLysIleMetIleThrAspPheGlyLeu 167  
Db 493 AACCTGCTGACTCACCCCTGAGGAGAACTCCAAGATAATGATCATTGTCGCTA 434  
Qy 168 SerLysMetGluClnAnGlyIleMetSerThrAlaCysGlyThrProGlyTyrValAla 187  
Db 433 TCCAAGATGGAGCAATAGAGTCATGTCCACAGCTGTGGGACCCAGGCTATGTGGCT 374  
Qy 188 ProGluValLeuAlaGlnLysProTyrSerLysAlaValAspCysTrpSerIleGlyVal 207  
Db 373 CCAGAAGTGTGCGCCAGAGAGCCCTACAGTAAGGCTGTGGACTGCTGCTCCATTGCTGTC 314  
Qy 208 IleThrTyrIleLeuLeuCysGlyTyrProPropheTyrGluGluThrGluSerLysLeu 227  
Db 313 ATCACATACATACTGCTGTGTGGGTATCCCTTTCTATGAAGAAACAGAAATCAAGGCTT 254  
Qy 228 PheGluLysIleLysGluGlyTyrTyrGluPheGluSerPropheTrpAspAspIleSer 247

Db 253 TTTGAGAAGATCAAGAGGTTACTAGAGTTTGAGTCTCCATTCTGGGATGACATTCT 194  
Qy 248 GluSerAlaLysAspPheIleCysHisLeuLeuGluLysAspProAnGluArGtYrThr 267  
Db 193 GAGTCAGCCCAAGACTTATTTGCCATCTCTTGAGAGAGGCCCAATGACATACACC 134  
Qy 268 CysGluLysAlaLeuSerHisProTrpIleAspGlyAsnThrAlaLeuHisArgAspIle 287  
Db 133 TGGGAGAAAGCCCTCAGACACACCTTGATTTGACGGGAACACAGCCCTGCACCGGACATC 74  
Qy 288 TyrProSerValSerLeuGlnIleGlnLysAsnPheAlaLysSerLysTrpArGlnAla 307  
Db 73 TACCATCTCTCAGCCCTCCAGATTTCAGAGAACTTTGCCAAGAGCAAGTGGGAGGCAAGCC 14  
Qy 308 PheAsnAlaAla 311  
Db 13 TTCATGCGGCC 2  
RESULT 15  
AL556476  
LOCUS AL556476 LTI\_NFL006\_PL2 962 bp mRNA linear EST 16-FEB-2001  
DEFINITION AL556476 LTI\_NFL006\_PL2 Homo sapiens cDNA clone CS0DK004YH19 5  
prime, mRNA sequence.  
ACCESSION AL556476  
VERSION AL556476.1 GI:12899184  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 962)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
source

Location/Qualifiers  
1..962  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DK004YH19"  
/clone\_lib="LTI\_NFL006\_PL2"  
/tissue\_type="placenta"  
/note="Vector: pCMVSPORT 6; Site\_1: Not1; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-stranded cDNA was digested with Not I and  
cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
vector. Library was normalized. Library was constructed by  
Life Technologies. Contact : Feng Liang Life Technologies,  
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com"

BASE COUNT 218 a 263 c 293 g 185 t  
ORIGIN

## Alignment Scores:

Pred. No.: 2.66e-88 Length: 962  
Score: 966.50 Matches: 189  
Percent Similarity: 83.46% Conservative: 38  
Best Local Similarity: 69.49% Mismatches: 42  
Query Match: 38.46% Indels: 4  
DB: 9 Gaps: 3

US-09-960-643-2 (1-476) x AL556476 (1-962)

Qy 12 TrpLysLysGlnThrThrAsnIleArgLysThrPheIlePheMetGluValLeuGlySer 31  
Db 150 TGG---AAGCAGCGGAGGACATTTAGACATCTACGACTTCCGAGATGTTCTGGGACG 206

```
Qy 32 GlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGlyLysLeuPheAlaLeu 51
Db 207 GGGGCCCTTCGCGAGGTGATCCTGGCAGAAGATAAGAGCAGCAGAGAGCTGGTGCCATC 266
Qy 52 LysCysIle--LysLysSerProAlaPheArgAspSerSerLeuGluAsnGluIleAla 70
Db 267 AATATGATTCACCAAGAGGCCCTGGAGGCCAAGGAGGAGCAGCATGGAGAAATGAGATTGCT 326
Qy 71 ValLeuLysLysIleLysHisGluAsnIleValThrLeuGluAspIleTyrGluSerThr 90
Db 327 GTCTGTCACAAGATCAAGACACCCCAACATTTGAGCCCTGGATGACATCTATGAGAGTGGG 386
Qy 91 ThrHisTyrThrLeuValMetGlnLeuValSerGlyGlyLeuPheAspArgIleLeu 110
Db 387 GGCACCTCTACCTCATCATGAGCTGGTGGTGGGGAGGAGCTCTTGACCGGTATTGTG 446
Qy 111 GluArgGlyValTyrThrGluLysAspAlaSerLeuValIleGlnGlnValLeuSerAla 130
Db 447 GAAAAAGGCTTCTACACGGAGCGGACGCCAGCGGCTCATCTCCAGGTGCTGGATGCT 506
Qy 131 ValLysTyrLeuHisGluAsnGlyIleValHisArgAspLeuLysProGluAsnLeuLeu 150
Db 507 GTGAATACCTGCTGATGACCTGGGCAATGTACACCGGGATCTCAAGCCAGAGAATCTGCTG 566
Qy 151 TyrLeuThrProGluAsnSerLysIleMetIleThrAspPheGlyLeuSerLysMet 170
Db 567 TACTACAGCTGGATGAAGACTCCAANAATCATGATCTCCGACTTTGGCCCTCTCCAAGATG 626
Qy 171 GluGlnAsnGly---IleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGlu 189
Db 627 GAGGACCCGGCAGTGTCCTCACCGCTGTGGAACTCCGGGATACGTGGCCCTGAA 686
Qy 190 ValLeuAlaGlnLysProTyrSerLysAlaValAspCysTrpSerIleGlyValIleThr 209
Db 687 GTCTGGCCCAAGAGCCCTACAGCAAGGCTGTGGATTGGTGTCCATAGGTGTCTATCGCC 746
Qy 210 TyrIleLeuLeuCysGlyTyrProPheTyrGluGluThrGluSerLysLeuPheGlu 229
Db 747 TACATCTTGCTTGGCGTTACCCYCCCTTCTATGACGAGAAATGATGCCAAACTCTTGAA 806
Qy 230 LysIleLysGluGlyTyrTyrGluPheGluSerProPheTrpAspIleSerGluSer 249
Db 807 CAGATTTGAAGGCCGAGTACGAGTTTGACTCTCCTTACTGGGACGACATCTCTGACTCT 866
Qy 250 AlaLysAspPheIleCysHisLeuLeuLysAspProAsnGluArgTyrThrCysGlu 269
Db 867 GCCAAGATTTCATCCGGCACTTKATGGAGAGGA-CCAGAGAAAAGATTACCTGTGAG 925
Qy 270 LysAlaLeuSerHisProTrpIleAspGlyAsnThr 281
Db 926 CAGGCCTTGCAGCACCCCATGGATTGCAGGAGATACA 961
```

Search completed: March 15, 2003, 06:39:01  
Job time : 2231 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 15, 2003, 05:09:07 ; Search time 113 Seconds  
(without alignments)  
2956.529 Million cell updates/sec

Title: US-09-960-643-2  
Perfect score: 2513  
Sequence: 1 MGRKEDDCSWKKQTTNIR.....VKASGSSHRAGQTGVCLIM 476

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 501302 seqs, 350932545 residues  
Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US09960643/runat\_07032003\_090518\_19918/app\_query.fasta\_1.647  
-DB=PublishedApplications NA -OFMT=fastap -SUEFTX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40 cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USRF=US09960643 @cgn1\_1\_57 @runat\_07032003\_090518\_19918  
-NCPU=6 -ICPU=3 -NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA.\*  
1: /cgn2.6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2.6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2.6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2.6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2.6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2.6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2.6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2.6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2.6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2.6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2.6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2.6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2.6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2.6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2513	100.0	1738	9 US-09-935-464-4	Sequence 4, Appli
2	2362	94.0	1383	9 US-09-935-464-2	Sequence 2, Appli
3	1250.5	49.8	1772	9 US-10-024-036B-1	Sequence 1, Appli
4	1246	49.6	1074	9 US-10-024-036B-3	Sequence 3, Appli

5	1237	49.2	1578	10 US-09-835-788A-6	Sequence 6, Appli
6	1029	40.9	1372	10 US-09-817-181-1	Sequence 1, Appli
7	865	34.4	157875	9 US-09-935-464-1	Sequence 1, Appli
8	760	30.2	480	9 US-09-935-464-6	Sequence 46, Appli
9	747.5	29.7	1503	10 US-09-797-039-3	Sequence 3, Appli
10	747.5	29.7	2297	10 US-09-797-039-1	Sequence 1, Appli
11	747.5	29.7	3124	12 US-10-153-921-1	Sequence 1, Appli
12	667	26.5	467	9 US-09-935-464-47	Sequence 47, Appli
13	645	25.7	470	9 US-09-935-464-48	Sequence 48, Appli
14	639.5	25.4	3552	9 US-09-934-406-1	Sequence 1, Appli
15	639.5	25.2	474	9 US-09-796-692-7777	Sequence 7777, Ap
16	625.5	24.9	2298	9 US-10-217-357-3	Sequence 3, Appli
17	625.5	24.9	2298	10 US-09-975-326-3	Sequence 3, Appli
18	625.5	24.9	2301	9 US-09-934-406-3	Sequence 3, Appli
19	625.5	24.9	2301	9 US-10-217-357-1	Sequence 1, Appli
20	625.5	24.9	2301	10 US-09-975-326-1	Sequence 1, Appli
21	611.5	24.3	7626	9 US-10-001-835-82	Sequence 82, Appli
22	606.5	24.1	1158	9 US-10-116-332-1	Sequence 1, Appli
23	591.5	23.5	5637	10 US-09-917-800A-1537	Sequence 1537, Ap
24	591	23.5	2824	10 US-09-992-481-5	Sequence 5, Appli
25	587	23.4	3579	9 US-10-024-036B-4	Sequence 4, Appli
26	585.5	23.3	1947	9 US-10-024-036B-6	Sequence 6, Appli
27	585.5	23.3	2454	10 US-09-992-481-3	Sequence 3, Appli
28	584.5	23.3	3705	12 US-10-044-090-227	Sequence 227, App
29	583.5	23.2	1808	10 US-09-740-627-2	Sequence 2, Appli
30	574.5	22.9	2061	12 US-10-096-960-1	Sequence 1, Appli
31	563.5	22.4	1473	9 US-09-938-842A-786	Sequence 786, App
32	562.5	22.4	2735	9 US-10-081-119-15	Sequence 15, Appli
33	556.5	22.1	1833	9 US-09-938-842A-2334	Sequence 2334, Ap
34	549.5	21.9	3061	10 US-09-880-107-2146	Sequence 2146, Ap
35	543	21.6	1488	9 US-09-938-842A-704	Sequence 704, App
36	539.5	21.5	1551	9 US-09-925-299-210	Sequence 210, App
37	539.5	21.5	1551	10 US-09-925-299-210	Sequence 210, App
38	537.5	21.4	1545	10 US-09-922-138-6	Sequence 6, Appli
39	537.5	21.4	1545	10 US-09-841-683-1	Sequence 1, Appli
40	537.5	21.4	2001	10 US-09-841-683-3	Sequence 3, Appli
41	537.5	21.4	2456	10 US-09-922-138-4	Sequence 4, Appli
42	528.5	21.0	1635	9 US-09-938-842A-2344	Sequence 2344, Ap
43	526.5	21.0	1800	9 US-09-938-842A-1063	Sequence 1063, Ap
44	526	20.9	1638	9 US-09-938-842A-1859	Sequence 1859, Ap
45	521	20.7	5926	10 US-09-969-708-302	Sequence 302, App

ALIGNMENTS

RESULT 1  
US-09-935-464-4  
; Sequence 4, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCH  
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300  
; PRIOR FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 1738  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-935-464-4

Alignment Scores:  
Pred. No.: 6,89e-232 Length: 1738  
Score: 2513.00 Matches: 476  
Percent Similarity: 100.00% Conservatlv: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0
US-09-960-643-2 (1-476) x US-09-935-464-4 (1-1738)			
Qy	1	MetGlyArgLysGluUAspAspCysSerSerTrpLysLysGlnThrThrAsnIleArg	20
Db	18	ATGGGTGAAAGGAAGAGATGACTGCAGTTTCCTGGGAAGAACAGACCAACATCCGG	77
Qy	21	LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal	40
Db	78	AAACCTTCATTTTATGGAAGTGTCTGGATCAGGAGCTTTCACAGATTTTCCTGGTG	137
Qy	41	LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe	60
Db	138	AAGCAAGACTGACTGGGAAGCTTTTGCTCTGAAGTGCATCAAGAAAGTCACCTGCCTC	197
Qy	61	ArgAspSerSerLeuGluUAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle	80
Db	198	CGGGACAGCGCTGGAGAATGAGATTGCTGTGTGAAAGATCAAGCATGAAACATT	257
Qy	81	ValThrLeuGluUAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal	100
Db	258	GTGACCTTGGAGGACATCTATGAGACACACCCACTACTACCTGGTCATGCAGCTTGT	317
Qy	101	SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla	120
Db	318	TCTGTGGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAGAAGATGCC	377
Qy	121	SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal	140
Db	378	AGTCGTGTGNTCCAGCAGCTCTTGTGGCAGTGAATACCTACATGAGATGGCATCGTC	437
Qy	141	HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluUAsnSerLysIle	160
Db	438	CACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCTCCCTGAAGAGAACTCTAAGATC	497
Qy	161	MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys	180
Db	498	ATGATCACTGACTTTGGTCTCTCCAAGATGAACAGAAATGCATGTCCACTGCCTGT	557
Qy	181	GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal	200
Db	558	GGGACCCAGGCTAGTGGCTCCAGAGTGTGGCCCAAGAACCTACAGCAAGCTGTG	617
Qy	201	AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyr	220
Db	618	GATTGCTGGTCCATCGGCGTCATCACCTACATATTGTCTGTGGATACCCCCCATTTCTAT	677
Qy	221	GluGluThrGluSerLysPheGluLysIleLysGlyTyrTyrGluPheGluSer	240
Db	678	GAAGAAACGGAGCTTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGATTTTGAGTCT	737
Qy	241	ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys	260
Db	738	CCATTCTGGATGACATTTCTGAGTCAGCCCAAGGACTTTATTGCCACTTGTCTTGAGAAG	797
Qy	261	AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn	280
Db	798	GATCCGAACGACGGGTACACCTGTGAGAAGCCCTTGATGTCATCCCTGGATTGACGGAAC	857
Qy	281	ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla	300
Db	858	ACAGCCCTCCACCGGACATCTACCCATCAGTCAGCTCCAGATCCAGAGAATTTTGCT	917
Qy	301	LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu	320
Db	918	AAGAGCAAGTGGAGCAAGCCTTCAACGCAGCAGCTGTGGTGCACCACCATGAGGAAGCTA	977
Qy	321	HisMetAsnLeuHisSerProGlyValArgProGluValGluUAsnArgProProGluThr	340
Db	978	CACATGAACCTGCACAGCCCGGGCGTCCGCCAGAGGTGGAGAACAGGCCCGCTGAAACT	1037

Qy	341	GlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaProVal	360
Db	1038	CAAGCCTCAAAACCTCTAGACCCAGCTCCCTGAGATCACCATCACCAGGACCTGTC	1097
Qy	361	LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg	380
Db	1098	CTGACACACAGTGTAGACTCCTCGCCCTGACCCATTACCTGCCAGCATGGCCGCCG	1157
Qy	381	ProThrAlaProGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer	400
Db	1158	CCCACCTGCCCTGGTGGCAGTCCCTCAACTGCTGCTCAATGGCTCCCTCCACATCAGC	1217
Qy	401	SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysCysSer	420
Db	1218	AGCAGCTGGTGCCCATGATCAGGGGTCCCTGGCCGCCGGCCCTGGCTGCTCTCC	1277
Qy	421	SerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCysSerGluProThrLeuLeu	440
Db	1278	AGCTGCTGNACATTGGGAGCAAGGAAGTCTCTCTACTGCTCTGAGCCACACTCCTC	1337
Qy	441	LysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAlaSer	460
Db	1338	AAAAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTCTATGTTACAGTTAAAGCCAGT	1397
Qy	461	GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet	476
Db	1398	GGCAGCTCCCACTGCCGGCAGGCAGACTGGAGTCTGTCTATTATG	1445
RESULT 2			
US-09-935-464-2			
; Sequence 2, Application US/09935464			
; Publication No. US20030027153A1			
; GENERAL INFORMATION:			
; APPLICANT: Meyer, Joanne			
; APPLICANT: Barrington-Martin, Rory			
; APPLICANT: Parker, Alexander			
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCH			
; FILE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA			
; FILE REFERENCE: 3322/11702 US1			
; CURRENT APPLICATION NUMBER: US/09/935,464			
; CURRENT FILING DATE: 2001-08-23			
; PRIOR APPLICATION NUMBER: US 09/757,300			
; PRIOR FILING DATE: 2001-01-09			
; NUMBER OF SEQ ID NOS: 90			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 2			
; LENGTH: 1383			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-935-464-2			
Alignment Scores:			
Pred. No.:	1,55e-217	Length:	1383
Score:	2362.00	Matches:	447
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	93.99%	Indels:	0
DB:	9	Gaps:	0
US-09-960-643-2 (1-476) x US-09-935-464-2 (1-1383)			
Qy	1	MetGlyArgLysGluUAspAspCysSerSerTrpLysLysGlnThrThrAsnIleArg	20
Db	1	ATGGGTGAAAGGAAGAGATGACTGCAGTTTCCTGGGAAGAACAGACCAACATCCGG	60
Qy	21	LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal	40
Db	61	AAACCTTTCATTTTATGGAAGTGTCTGGATCAGGAGCTTTCACAGATTTTCCTGGTG	120
Qy	41	LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe	60
Db	121	AAGCAAGACTGACTGGGAAGCTTTTGCTCTGAAGTGCATCAAGAAAGTCACCTGCCTC	180



```
Db 289 GAAGAGAAGCAACTGGCAAGCTCTTTGCTGTGAAGTGTATCCCTAAGAAGCGCTGAAG 348
Qy 60 PheArgSerSerLeuGluAsnGluLeuAlaValLeuLysLysLeuHisGluAsn 79
Db 349 GGCACGAAGAAGCAGATAGATAGATAGCGCTCGTGAAGAAGATTAAAGCATGAATAAT 408
Qy 80 IleValThrLeuGluAspIleThrHisGluSerThrHisValLeuValMetGlnLeu 99
Db 409 ATTGTTGCCCTGGAAGACATTTATGAAGCCCAATCACTGACTGCTGTGTCATGCAGCTG 468
Qy 100 ValSerGlyGlyLeuLeuPheAspArgIleLeuGluArgGlyValThrGluLysasp 119
Db 469 GTCTCCGGTGGAGAGCTGTTGACCGGATAGTGGAGAGGGTTTTATACAGAGAAGAT 528
Qy 120 AlaSerLeuValIleGlnValLeuSerAlaValLysThrLeuHisGluAsnGlyIle 139
Db 529 GCAGACACTCTGATCCGCCAAGTCTTGGACGGGTGCTACTCTCCACAGAATGGGCATC 588
Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuThrProGluLysSerLys 159
Db 589 GTCCACAGAGACCTCAAGCCGCAAAATCTTGTACTACAGTCAAGATGAGGAGTCCAAA 648
Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178
Db 649 ATAATGATCAGTGACTTTGGATTGTCAAAAATGGAGGCCAAAGGAGATGTGATGTCCACT 708
Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198
Db 709 GCCTGTGGAACTCCAGGCTATGTGCTCTGAAAGTCTGCCCGCCAGAAACCTTACAGCAA 768
Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
Db 769 GCGGTTGACTCGTGGTCCATCGAGTGAATGCTTACTATCTGCTGCGCTACCTCTCT 828
Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238
Db 829 TTTTATGATGAANAATCAAGCTCTTTGAGCAGATCCTCAAGCGGGAATATGAGTTT 888
Qy 239 GluSerProPheTrpAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
Db 889 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAGACTTCAATTCGGAACCTGATG 948
Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleasp 278
Db 949 GAGAAGGACCCGAAATAAAGATACACGTGTGACAGCAGCTCGGCACCCATCGATCGCT 1008
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
Db 1009 GGTGACACAGCCCTCAACAAAACATCCAGAGTCCGTGAGCGCCAGATCCGGAAAAC 1068
Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318
Db 1069 TTTGCCAAGAGCAATGGAGACAAGCATTTAATGCCACGCGCTCGTGAGACATATGAGA 1128
Qy 319 LysLeuHisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArg----- 336
Db 1129 AAACATACCTCGCGCAGCAGCCCTGGACAGTTCAAATGCAAGTGTTCGAGCAGCCCTCAGT 1188
Qy 337 ---ProProGluThrGlnAlaSerGluThrSerArgProSerSerProGluIleThrIle 355
Db 1189 TTGCCAGCCAAAAGACTGTGGTATGTAGCAAAACAGATC-CCTCAGCTGACACTG 1247
Qy 356 ThrGluAlaProValLeuAspHisSerValAlaLeuProAla----- 369
Db 1248 AA-GACGAGCTGGGTGGAGAGGAGCGGCATCTGCGGAGCAGCCTCCTGTTTGGCC 1306
Qy 370 -----LeuThrGlnLeuProCys----- 375
Db 1307 AGCGGCTTTCTATACTAATATCCATGTGTCATGCGACCCCTAGGACTTTTTTAACATGTAAT 1366
Qy 376 -----GlnHis---GlyArg----- 379
Db 1367 CACTGGGCGGGTGCAGTGGCTCACGCTCTAATCCCAACACTTCTGGGAGGCTGAGGCAG 1426
```

```
Qy 380 -----ArgProThrAlaPro----- 384
Db 1427 GAGGACTGTTTGTAGTTCAGGAGTTTTTAAGACCACTGACCAACATGCTGAAACCCCATC 1486
Qy 385 -----GlyGlyArgSerLeuAsnCysLeuValAsn 394
Db 1487 TCTACTAAAATATAAAATTAAGCGGGTGTGGTGGCAGCACCTGTAATGTACGTACTT 1546
Qy 395 GlySerLeuHisIleSerSerSerLeuValProMetHisGlnGlySerLeuAlaAlaGly 414
Db 1547 GGGAGGCTGAGGCGAGGAGAACTTGAACCA-----GGA 1582
Qy 415 ProCysGlyCysCysSerSerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCys 434
Db 1583 AGCGGAGGTTGC-----AATGAGCTGAGATCACCACCTGC 1618
Qy 435 SerGlu-----ProThrLeuLeuLysLysAlaAsnLysLysGlnAsn 448
Db 1619 ACTCCAGCTGGGTGACAGATTGAGACTCCCTCTCAAAAAAAGGAAATCATTTGAAC 1678
Qy 449 PheLysSerGluValMetValProValLysAlaSerGly 461
Db 1679 ACTCGTGAACCTAGGTATTGCTATTTCCATTTACGGT 1717
RESULT 4
US-10-024-036B-3
; Sequence 3, Application US/10024036B
; Publication No. US20030028004A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 68730 and 69112, Protein Kinase
; TITLE OF INVENTION: Molecules and Uses Therefor
; FILE REFERENCE: MPI2000-521PIR(M)
; CURRENT APPLICATION NUMBER: US/10/024,036B
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/258222
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1074
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-024-036B-3
Alignment Scores:
Pred. No.: 1.59e-110 Length: 1074
Score: 1246.00 Matches: 237
Percent Similarity: 85.49% Conservative: 40
Best Local Similarity: 73.15% Mismatches: 45
Query Match: 49.58% Indels: 2
DB: Gaps: 2
US-09-960-643-2 (1-476) x US-10-024-036B-3 (1-1074)
Qy 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
Db 1 ATGGCCCGGAGAGACGGCGAGAGCAGCTCCTCTGAAAAAGCAAGCTGGAAGACATCAAG 60
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 61 AGATCTTCGAGTCTCAAGAGACCCCTCGAACCAGGGGCTTTTCGGAAGTGGTTTACGT 120
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59
Db 121 GAAAGAAAGCACTGGCAAGCTCTTGTCTGGAAGTGTATCCCTAAGAGCGCTGAAG 180
Qy 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
Db 181 GCAAGGAAGCAGCATAGAGATGAGATGCGCTCTGAGAAAGATTAAAGCATGAAAT 240
Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99
```

```
Db 241 ATTGTTCCCTGGGAAGACATTTATGAAGCCCAAAATCACTGTACTGTGGTCATGCAGCTG 300
Qy 100 ValSerGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119
Db 301 GTGTCGCGTGGAGAGCTGTTGACCGGATAGTGGAGAGGGGTTTATACAGAGAAGAT 360
Qy 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139
Db 361 GCCAGCACTCTGATCCGCCAAGCTTTGGAGCGCGTGTACTATCTCCACAGAAATGGGCATC 420
Qy 140 ValHisArgAspLeuLysProGluAsnLeuTyrLeuThrProGluGluAsnSerLys 159
Db 421 GTCCACAGACCTCAAGCCCGAAATCTCTGTACTACAGTCAAGATGAGGAGTCCAAA 480
Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178
Db 481 ATAATGATCAGTGACTTTGGATTGTCAAAAATGGAGGGCAAGAGATGTGATGTCCACT 540
Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198
Db 541 GCCTGTGGAACTCCAGCGTATGCGCTCCCTGAAGTCTCCGCCAGAAACCTTACAGCAA 600
Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
Db 601 GCCGTTGACTGCTGCTCCATCGGAGTGATGCTCATACATCTTGCTCTCGCGCTACCTCCT 660
Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleGlyTyrTyrGluPhe 238
Db 661 TTTTATGATGAAATGACTCAAGCTCTTTTGAGCAGATCTCTCAAGGCGGAATATGAGTTT 720
Qy 239 GluSerProPheTrpAspPheSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
Db 721 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAGACTTCATTCGGAACCTGATG 780
Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278
Db 781 GAGAAGGACCGCAATAAAGATACACGTGTGACGAGCAGCTCGCGCACCCATGGATCGCT 840
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
Db 841 GGTGACACAGCCCTCAACAAAAACATCCACGAGTCCGTCAGCGCCAGATCCGGAAAAAC 900
Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318
Db 901 TTTCCCAAGAGCAATGGACACAGCATTTAATGCCACGCCGCTGTGAGACATATGAGA 960
Qy 319 LysLeuHisMet 322
Db 961 AAACCTACACCTC 972
```

## RESULT 5

```
US-09-835-788A-6
; Sequence 6, Application US/09835788A
; Patent No. US20020077458A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides, a
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: PTO18P1
; CURRENT APPLICATION NUMBER: US/09/835,788A
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US00/28666
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/159,585
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/167,246
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1578
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-835-788A-6
```

## Alignment Scores:

```
Pred. No.: 2,07e-109 Length: 1578
Score: 1237.00 Matches: 235
Percent Similarity: 86.16% Conservative: 39
Best Local Similarity: 73.90% Mismatches: 42
Query Match: 49.22% Indels: 2
DB: 10 Gaps: 2
```

US-09-960-643-2 (1-476) x US-09-835-788A-6 (1-1578)

```
Qy 7 AspAspCysSerSerTrpLysGlnThrThrAsnIleArgLysThrPheIlePheMet 26
Db 15 GAGAGCAGCTCTCTCTGAAAAAGCAGCTGAAGACATCAAGAAAGATCTTCGAGTTCAAA 74
Qy 27 GluValLeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGly 46
Db 75 GAGACCTTCGGNACCGGGGCGCTTTCCGAGTGGTTTAGCTGAAGAGAAGCAACTGGC 134
Qy 47 LysLeuPheAlaLeuLysCysIle---LysLysSerProAlaPheArgAspSerSerLeu 65
Db 135 AAGCTCTTTGCTGTGAAGTGTATCCTTAAGAAGCGCTGAAGGCAAGAAAGCAGCAT 194
Qy 56 GluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIleValThrLeuGluAsp 85
Db 195 GAGATGAGATAGCGCTCTGAGAAAGATTAAAGCATGAAATATTTGTCCTCGAAGAC 254
Qy 86 IleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuValSerGlyGlyGluLeu 105
Db 255 ATTTATGAAGCCCAAAATCACCTGTACTTGGTCATGAGCTGGTGTCCGGTGGAGAGCTG 314
Qy 106 PheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAlaSerLeuValIleGln 125
Db 315 TTTGACCGGATAGTGGAGAGGGGTTTATACAGAGAAGGATGCGACACTCTGTATCGCG 374
Qy 126 GlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleValHisArgAspLeuLys 145
Db 375 CAAGTCTTGGAGCGGTGTACTATCTCCACAGAATGGGCATCGTCCACAGAGACTCAAG 434
Qy 146 ProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIleMetIleThrAspPhe 165
Db 435 CCCGAAATCTCTTCTACTACAGTCAAGATGAGGAGTCCAAAATAATGATCAGTGACTTT 494
Qy 166 GlyLeuSerLysMetGluGlnAsnGly---IleMetSerThrAlaCysGlyThrProGly 184
Db 495 GGATTTGTCAAAAATGGAGGGCAAGAGATGTGATGTCCACTGCTGTGGAACTCCAGGC 554
Qy 185 TyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaValAspCysTrpSer 204
Db 555 TATGTCGCTCTCGAAGTCTCTGCCCGAGAAACCTTACAGCAAGAGCGGTGACTGTGTGTC 614
Qy 205 IleGlyValIleThrTyrIleLeuLeuCysGlyTyrProProPheTyrGluGluThrGlu 224
Db 615 ATCGGAGTGATTGCTCTACATCTTGCTCTGCGCTACCTCTCTTTTATGATGATAAATGAC 674
Qy 225 SerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSerProPheTrpAsp 244
Db 675 TCCAAGCTCTTTTGAGCAGATCTCTCAAGGCGGAATATGAGTTTGACTCTCTCCCTACTGG 734
Qy 245 AspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLysAspProAsnGlu 264
Db 735 GACATCTCCGACTCTGCAAAAAGACTTCATTCGGAACCTGATGGAGAGAGCCCGAAATAA 794
Qy 265 ArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsnThrAlaLeuHis 284
Db 795 AGATACAGTGTGAGCAGGAGCTCGGCACCCATGGATCGCTGTGACACAGCCCTCAAC 854
Qy 285 ArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAlaLysSerLysTrp 304
Db 855 AAAACATCCACGAGTCCGTCAGGCCCCAGATCGGAAAAAACTTTGGCAAGAGCAAAATGG 914
```

QY 305 ArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeuHisMet 322  
|||||  
Db 915 AGACAAGCATTTAATGCCAGCCGCTGTCGACATATGAGAAACTACACCTC 968

## RESULT 6

US-09-817-181-1  
; Sequence 1, Application US/09817181  
; Patent No. US2002014247A1  
; GENERAL INFORMATION:  
; APPLICANT: MERKULOV, Gennady et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001189  
; CURRENT APPLICATION NUMBER: US/09/817,181  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1372  
; TYPE: DNA  
; ORGANISM: Human  
US-09-817-181-1

Alignment Scores:  
Pred. No.: 1.59e-89 Length: 1372  
Score: 1029.00 Matches: 224  
Percent Similarity: 64.20% Conservativity: 45  
Best Local Similarity: 53.46% Mismatches: 90  
Query Match: 40.95% Indels: 60  
DB: 10 Gaps: 9

US-09-960-643-2 (1-476) x US-09-817-181-1 (1-1372)

QY 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79  
|||||  
Db 198 TTCAGAT-----GTTCTGGGCAGCATCAAGACCCCAAC 233  
QY 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99  
|||||  
Db 234 ATTGAGCCCTGGATGACATCTATGAGAGTGGGGCCACCTCTACCTCATCATGAGCTG 293  
QY 100 ValSerGlyGlyLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119  
|||||  
Db 294 GTGTGGGTGGGGAGCTCTTTCACCGTATGTGGAAGAGCTTCTACACGGAGCGGAC 353  
QY 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139  
|||||  
Db 354 GCCAGCGGCTCATCTCCAGTGTGGATGCTGTGAATACCTGCATGACCTGGGCATT 413  
QY 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159  
|||||  
Db 414 GTACACCGGGATCTCAAGCAGAGAATCTGTGTACTACAGCTGGATGAAGACTCCAAA 473  
QY 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly--IleMetSerThr 178  
|||||  
Db 474 ATCATGATCTCCGACTTGGCTCTCCAGATGGAGACCGCGGAGTGTCTCTCCACC 533  
QY 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198  
|||||  
Db 534 GCCTGTGGAATCCGGGATACCTGGCCCTGAAGTCTGGCCCAAGCCCTACAGCAAG 593  
QY 199 AlaValAspCysTrpSerIleGlyValIleThrThrIleLeuLeuCysGlyTyrProPro 218  
|||||  
Db 594 GCTGTGGATTGTGTCTCCATAGTGTCTATCGCTACATCTTGTCTGGGGTTACCTCCC 653  
QY 219 PheTyrGluGluThrCysLeuSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238  
|||||  
Db 654 TTCTATGACGAGAATGATGCCAACTCTTTGAACAGATTTGAAGGCGGAGTACAGATT 713  
QY 239 GluSerProPheTrpAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258  
|||||  
Db 714 GACTCTCCTTACTGGGACGACATCTCTGACTCTGCCAAGATTTCATCCGGCATTGATG 773

QY 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTyrPileAsp 278  
|||||  
Db 774 GAGAAGGACCCAGAGAAAGATTACCTGTGACAGSCCTTGACAGCACCCTGATTGCA 833  
QY 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298  
|||||  
Db 834 GGAGATACAGCTCTAGATAAGAATATCCACAGTCGGTGTAGTGAGCAGATCAAGAGAAC 893  
QY 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArg 318  
|||||  
Db 894 TTTCGCAAGAGCAGTGAAGCAGCTTCAATGCCACGGCTGTGTGTCGCGCACATGAGG 953  
QY 319 LysLeuHisMetAsnLeuHisSerProGly-----Val 329  
|||||  
Db 954 AAACCTGACGTGGGCACACGAGGAGGCGAGGGCAGCGGAGCCATGGGAGCTG 1013  
QY 330 ArgProGluValGluAsnArgProProGluThrGlnAlaSerGluThrSerArgProSer 349  
|||||  
Db 1014 CTGACACCACTGGCTGGGGGGCG-----GCAGCTGGCTGTGCTCTCGAGACTGC 1064  
QY 350 SerProGluIleThrIleThrGluAlaProValLeuAspHisSerValAlaLeuProAla 369  
|||||  
Db 1065 TGCCTGAGCGCGGCACAGAACTGTCCCCACACTGCCCCACCACTC----- 1112  
QY 370 LeuThrGlnLeuProCysGlnHisGlyArgArgProThrAlaProGlyGlyArgSerLeu 389  
|||||  
Db 1113 -----TAGGGCCCTGGACCTCGGTCATG 1136  
QY 390 AsnCysLeuValAsnGlySerLeuHisIleSerSerSerLeuValProMetHis----- 407  
Db 1137 ATCTCTGCTGGGAGG-----CTTGGGGGACGCTGTCCCTTCCCTCCCTCG 1187  
QY 408 GlnGlySerLeuAlaAla----- 413  
Db 1188 ACCGGAGTTCTGCGCCTCTCCCTCACCTGCTTCCCTACCACTCCTCAGTGCAT 1247  
QY 414 GlyProCysGlyCysCysSerSerCysLeuAsnIleGlySerLysGlyLysSerSerTyr 433  
|||||  
Db 1248 TTTCCATACAAATGTTTCTATTTTATTGTTCTTCTTGTATTAAGGAAG----- 1298  
QY 434 CysSerGluProThrLeuLysAlaAsnLysGlnAsnPheLysSerGlu 452  
|||||  
Db 1299 ---ATAAACCATCTTAAAAA----- 1352

## RESULT 7

US-09-935-464-1  
; Sequence 1, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCH  
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300  
; PRIOR FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 157875  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-935-464-1

Alignment Scores:  
Pred. No.: 1.15e-70 Length: 157875  
Score: 865.00 Matches: 202  
Percent Similarity: 42.44% Conservativity: 0  
Best Local Similarity: 42.44% Mismatches: 3

Query Match: 34.42% Indels: 272  
DB: 9 Gaps: 2  
US-09-960-643-2 (1-476) x US-09-935-464-1 (1-157875)  
Qy 272 LeuSerHisProTrp-IleAspGlyAsnThrAlaLeuHisArgAspIleTyProSerVa 291  
Db 145909 CTCTACATCCACAGGATTGACGAAACACAGCCCTCCACCGGACATCATCCCATCAGT 145968  
Qy 291 lSerLeuGlnIleGlnLysAsnPheAlaLysSerLysTrpArg----- 305  
Db 145969 CAGCTCCAGATCCAGNAGAACTTGTCTAAGACCAAGTGGAG-GGTAAGCTGTCTCTCC 146027  
Qy 305 ----- 305  
Db 146028 AGGGGTGGGAAAGCTGTTCTGGGCCCTTGGAGGCTGGGCTGCGAGGGCTGACATNAGG 146087  
Qy 305 ----- 305  
Db 146088 GCTTTCCTTGGGATGTCCAGAAAGGCAATTGGAGCTCCGTATACCCCTCTCTGAAATGAGAA 146147  
Qy 305 ----- 305  
Db 146148 GTGGCACCCAGGTTTCAAGAGGCCACAGGCCAAAGGGAAGTTTAAAGCTCCCAAGGCCCT 146207  
Qy 306 -----GlnAlaPheAsnAlaAl 311  
Db 146208 CTGAGGTTGCAGAGGCCAGAGGCTGCTCTGTGTCTCTTAGCAAGCCCTTCAAGCGCAGC 146267  
Qy 311 aAlaValValHisHisMetArgLysLeuHisMetAsnLeuHisSerProGlyValArgPr 331  
Db 146268 AGCTGTGTGCACCATGAGGAGCTACACATGAACCTGCACAGCCCGGGGCTCCGCC 146327  
Qy 331 oGluValIleLysAsnArgProGluThrGlnAlaSerGluThrSerArgProSerSerPr 351  
Db 146328 AGAGGTGAGAACAGCGCGCTGAAACTCAAGCCTCAGAAACCTCTAGACCCAGCTCCCC 146387  
Qy 351 oGluIleThrIleThrGluAlaProValLeuAspHisSerValAlaLeuProAlaLeuTh 371  
Db 146388 TGAGATCACCATCACCAGGACACTGTCTGGACACAGTGTAGCACTCCCTGCCCTTGAC 146447  
Qy 371 rGlnLeuProCysGlnHisGlyArgArgProThrAlaProGlyArgSerLeuAsnCy 391  
Db 146448 CCAATTACCTTGCCAGCATGGCCGCGGCCACTGCCCTGGTGGCAGGTCCCTCAACATG 146507  
Qy 391 sLeuValAsnGlySerLeuHisIleSerSerSerLeuValProMetHisGlnGlySerLe 411  
Db 146508 CCTGGTCAATGGCTCCCTCCACATCAGCAGCAGCCTGGTGGCCATGCATCAGGGGTCCCT 146567  
Qy 411 uAlaAlaGlyProCysGlyCysSerSerCysLeuAsnIleGlySerLysGlyLysSe 431  
Db 146568 GGCCGCGGGCCCTGTGGCTGTCTGCTCCAGCTGCCTTGAACATTTGGAGGCAAGGAAGTC 146627  
Qy 431 rSerTyrCysSerGluProThrLeuLeuLysLysAlaAsnLysLysGln----- 447  
Db 146628 CTCCTACTGCTCTGAGCCACACTCCCTCAAAAGGCCAACAAAAACAGTACGTATTTT 146687  
Qy 447 ----- 447  
Db 146688 AGCCAAAGATGGAGCCCGCAGCTTGGGTCTGAAAGAAATCGGTCAACAGGACTGAAAGAAA 146747  
Qy 447 ----- 447  
Db 146748 TGGACACAAAGGCCTCTCCACTCATAGCAGCTATATAGGGAGGATGAGTTCGGATG 146807  
Qy 447 ----- 447  
Db 146808 AGGGGCAAGGAAAAATGCTTCCAGCCCTGTCCCATCATCTACTAGTGTGAGGCCATTTG 146867  
Qy 447 ----- 447  
Db 146868 ACGAGTCAATGCTCTGCTTCTATTTTGCTCATCTGTAAACTGGGTGAAGTCTATTATAC 146927

Qy 447 ----- 447  
Db 146928 CCAATTAGCCTGCTCAGAGAGGTACTATAAAAGTTAGTGAAGCAACAGAGATGATAAA 146987  
Qy 447 ----- 447  
Db 146988 AGCCCAAGAAAACACTGGGTATTGTGCAAGTGGTGCAGCATCAGATTCCCGAGTTTCCAGGG 147047  
Qy 447 ----- 447  
Db 147048 CCTCTGTGCTTTGCAGAAGCAAGTGTAGGGGGCTTGGTTATCTTTATCTTTTCTCAGG 147107  
Qy 447 ----- 447  
Db 147108 TTCTTGAACCTTCTGAACCTCATCAGGGGAGAGCTGAGTGGATGCCACAGGCACAGAGAA 147167  
Qy 447 ----- 447  
Db 147168 CTACCACCTCTGCCCTGCCCTGCCCTCGACTCTTCTTCCCTCAATACTTCAAGGTTTGTGC 147227  
Qy 448 -----AsnPhelYsSerGluValMetValProValLysAlaSerGlyS 462  
Db 147228 TTCATTTCTTCTAGGAAGTTCAAAGTCGGAGGTCTACCGATTAAAGCCAGTGGCA 147287  
Qy 462 erSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476  
Db 147288 GCTCCCACTGCCGGGCGAGGAGCAGACTGGAGTCTGTCTCATATG 147331  
RESULT 8  
US-09-935-464-46  
; Sequence 46, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSY  
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300  
; PRIOR FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 46  
; LENGTH: 480  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc-feature  
; OTHER INFORMATION: n-a or g or c or t/u  
US-09-935-464-46

Alignment Scores:  
Pred. No.: 2,22e-64 Length: 480  
Score: 760.00 Matches: 150  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 30.24% Indels: 0  
DB: 9 Gaps: 0

US-09-960-643-2 (1-476) x US-09-935-464-46 (1-480)

Qy 1 MetGlyArgLysGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20  
Db 30 ATGGGTCGAAAGGAGAAAGATGACTGCAGTTCCTGGAGAAACAGACACCAACATCCGG 89  
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
Db 90 AAAACCTTCATTTTATGGAAGTGTGGATCAGGAGCTTCTCAGAAGTTTCTCCTGGTG 149  
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60



```
QY 449 eLysSerGluValMetVal-----ProValLysAlaSerGlySerSerHisCysAr 466
Db 1356 GCCACCAAGCAGCTCCACCCCTGAGCGGCTATGGCCCGGAGCAGACAGCCCC 1415
QY 466 gAlaGlyGlnThrGly 471
Db 1416 AGAGGGCCGACAGGC 1431
RESULT 10
US-09-797-039-1
; Sequence 1, Application US/09797039
; Patent No. US20020042099A1
; GENERAL INFORMATION:
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 2504, 15977, AND 14760, NOVEL PROTEIN
; TITLE OF INVENTION: KINASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 10448-017001
; CURRENT APPLICATION NUMBER: US/09/797,039
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (154)...(1656)
; NAME/KEY: misc_feature
; LOCATION: (1)...(2297)
; OTHER INFORMATION: n = A,T,C or G
US-09-797-039-1
Alignment Scores:
Pred. No.: 3,74e-62 Length: 2297
Score: 747.50 Matches: 172
Percent Similarity: 53.29% Conservative: 87
Best Local Similarity: 35.39% Mismatches: 175
Query Match: 29.75% Indels: 53
DB: 10 Gaps: 9
US-09-960-643-2 (1-476) x US-09-797-039-1 (1-2297)
QY 15 GlnThrThrAsnIleArgLysThrPheIlePheMetGluValLeuGlySerGlyAlaPhe 34
Db 199 CAGCATCGGAGGTGACTGACAGATATGATTTGGGACAGGTATCAAGACTGAGGAGTTT 258
QY 35 SerGluValPheLeuValLysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle 54
Db 259 TGTGAATCTTCGGCGCAGGACAGACAGCAGGCAAGCTGCACACCTGCGAAGAGTTTC 318
QY 55 LysLysSerProAlaPheArgAsp-----SerSerLeuGluAsnGluIle 69
Db 319 CAGAAG-----CGGGACGCGCGCAAGGTGCGGAAAGCTGCCCAAGAACGAGATA 366
QY 70 AlaValLeuLysLysIleLysHisGluAsnIleValThrLeuLysAlaPheIleTyrGluSer 89
Db 367 GGCATCTCAAGATGTGTGAAGATCCCAATCTACAGCTGGTGATGTGTGTGACC 426
QY 90 ThrThrHisTyrTyrValMetGlnLeuValSerGlyGlyGluLeuPheAspArgIle 109
Db 427 CGCAAGGAGTACTTATCTTCCTGGAGCTGGCCAGGAGGAGGTGTTGACTGGATC 486
QY 110 LeuGluArgGlyValTyrThrGluLysAspAlaSerLeuValIleGlnValLeuSer 129
Db 487 CTGACACAGGGCTACTACTCGGAGCGAGACACAAGCAACCTGTGTACGGCAAGTCTCTGGAG 546
QY 130 AlaValLysTyrLeuHisGluAsnGlyIleValHisAspLeuLysProGluAsnLeu 149
Db 1569 AGAGGGCCGACAGGC 1584
```

```
Db 547 GCCGTGGCCTATTTCGCACTCACTCAAGATCGTCGACAGGAATCTCAAGCTGGAGAACTG 606
QY 150 LeuTyrLeuThrProGluGluAsnSerLysIleMetIleThrAspPheGlyLeuSerLys 169
Db 607 GTTATTACAAACCGGCTCAAGAACTCGAAGATTGTCATCAGTCACTTCCATCTGCCTAAG 666
QY 170 MetGluGlnAsnGlyIleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGlu 189
Db 667 CTAGAA--AATGGCCTCATCAAGGAGCCCTGTGGGACCTATCTGGCCCGCAGAG 723
QY 190 ValLeuAlaGlnLysProTyrSerLysAlaValAspCysTrpSerIleGlyValIleThr 209
Db 724 GTGGTAGCCCGGACGCGTATGGAGCCCTGTGACATGCTGGCCATTTGGACTCATG 783
QY 210 TyrIleLeuLeuCysGlyTyrProPheTyrGluGluThrGlu----- 224
Db 784 TACATCTCTGCTTTCAGGCAATCCACCTTCTATCAGGAGGTGGAAGAAGATGATTATGAG 843
QY 225 -----SerLysLeuPheGluLysIleLysGluGlyTyrGluPheGluSerPro 241
Db 844 AACCATGATAAGAAATCTCTCCGCAAGATCTCGTGTGCTGACTATGATTTGACTCTCCA 903
QY 242 PheTyrAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLysAsp 261
Db 904 TATTTGGGATGATTTTCGAGCGACCAAGACCTGTGTCACAGGCTGATGGAGTGGAG 963
QY 262 ProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTyrIleAspGlyAsnThr 281
Db 964 CAAGACCGCGGATCACTGCAAGAGAGGCCATCTCCCATGATGATTTCTGGCAATGCT 1023
QY 282 AlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAlaLys 301
Db 1024 GCTTCTGATAAGAACATCAAGGATGCTGTGTCGCCAGATGAAAGAACTTTGCCAGG 1083
QY 302 SerLysTyrArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeuHis 321
Db 1084 GCCAAGTGAAGAGGCTGTCGAGTGACACCTCTGAAACGGCTCCGGGACCAAGAG 1143
QY 322 MetAsnLeuHisSerProGly-ValArgProGluValGluAsnArgProProGlu----- 339
Db 1144 CAGTCCACGCGGTGCGACGCCAGTCCAGCCACAGACACTGCCACCCCGC-GGGGC 1202
QY 340 -----ThrGlnAlaSerGluThrSerArgProSerSerProGluIleThrI 355
Db 1203 TGCAGGTGGGGCCACAGCTGCGAGTGGCAGTGAGCTACCTCAGCCCTGAGGTGATGC 1262
QY 355 eThrGluAlaProValLeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCy 375
Db 1263 TGCTCGTCTGCTGCAAGAGTGTATAAT----- 1287
QY 375 sGlnHisGlyArgArgProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnG 395
Db 1288 -----GTGGCCCGCGAGCCGCTAGTGCCACCCAGCCAGCCAGATGG 1328
QY 395 ySerLeu-----HisIleSerSerLeuValProMetHisGlnGlySerLeuAlaAl 413
Db 1329 AAGTGCCACCCAGCCACTGATGCGAGTGCACCCAGCCAGCCAGCATCACTCC 1388
QY 413 aGlyProCysGlyCysCysSerSerCysLeuAsn-----lleGlySerLysG 429
Db 1389 AGCCACTGATGGAGTGTACCCCGCAGCCAGGAGCGCTACTCCAGCCACTGATG 1448
QY 429 yLysSerSerTyrCysSerGluProThrLeuLeuLysLysAlaAsnLysLysGlnAsnPh 449
Db 1449 GAGAGCCACACCGCCACAGAGAGACACTGTGCCACCACCCAAAGCAGTGCATGCT 1508
QY 449 eLysSerGluValMetVal-----ProValLysAlaSerGlySerHisCysAr 466
Db 1509 GGCCACCAAGCAGCTGCGACCCCTGTGACCGGCTATGGCCCGGAGCAGACAGCCCC 1568
QY 466 gAlaGlyGlnThrGly 471
Db 1569 AGAGGGCCGACAGGC 1584
```

**RESULT 11**

US-10-153-921-1  
 ; Sequence 1, Application US/10153921  
 ; Patent No. US20020142430A1  
**GENERAL INFORMATION:**  
 ; APPLICANT: YAN, Chunhua et al.  
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES THEREOF  
 ; FILE REFERENCE: CLOO612DIV  
 ; CURRENT APPLICATION NUMBER: US/10/153,921  
 ; PRIOR FILING DATE: 2002-05-24  
 ; PRIOR APPLICATION NUMBER: 60/207,281  
 ; PRIOR FILING DATE: 2000-05-30  
 ; PRIOR APPLICATION NUMBER: 09/734,030  
 ; PRIOR FILING DATE: 2000-12-12  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1:  
    LENGTH: 3124  
 ; TYPE: DNA  
 ; ORGANISM: HOMO SAPIEN

US-10-153-921-1

Alignment Scores:  
Pred. No.:                 5,95e-62          Length:                 3124  
Score:                     747.50           Matches:                 177  
Percent Similarity:       54.91%           Conservative:             80  
Best Local Similarity:    37.82%           Mismatches:            150  
Query Match:              29.75%          Indels:                 65  
DB:                        12             Gaps:                 10

US-09-960-643-2 (1-476) x US-10-153-921-1 (1-3124)

Qy	15	GlnThrThrAsnIleArgLysThrPheLePheMetGluValLeuGlySerGlyAlapHe	34
Db	256	CAGCCATCGGGAGGTACTGCACAGATAATTGTTGGCAGCATCTCAAGACTGAGGAGTTC	315
Qy	35	SerGluValPheLeuValLysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle	54
Db	316	TGTGAATCTTC CGGCCAACAGCACAGCAGCGAAAGTGCACACTGCAAGAAGTTC	375
Qy	55	LysLysSerProAlaPheArgASP-----SerSerLeuGluAsnGluIle	69
Db	376	CAGAAG-----CGGAGCGCGCGCAAGTGCGGAAAAGTGCCACAAGACGAGATA	423
Qy	70	AlaValLeuLysLysIleLysHisGluAsnIleValThrLeuGluAspIleTyrrGlusEr	89
Db	424	GGCATCTCAGATGGTGTAAGCATCCCCAACATCTCAGCTGGTGATGTGTTGTGACC	483
Qy	90	ThrThrIsTYRtyrLeuValMetGlnLeuValSerGlyGlyGluLeuPheAspargile	109
Db	484	CGCAAGGAGTAGTACTTATCTTCCTGGAGTGCCACGGGAGGAGGTGTTTGACTGGATC	543
Qy	110	LeuGluargGlyValTyrrThrGluLysaspAlasErLeuValIIeGInGluValser	129
Db	544	CTGACCAAGGCGCTACTACTCGGAGCGAGACACAAGCAACGTGGTACGCAAAGTCTCTGGAG	603
Qy	130	AlaValLystrYeuHISgluasnglytleValHisArgaspLeuLysProgluaSnLeu	149
Db	604	CCCGTGGCCATTTGGCACTCACTCAAGATCGTCACAGGAAATCTCAAGCTGGAGAACTG	663
Qy	150	LeutyrlEuThrProGluGluasnSerlysIlleMettlleThrAsppheGlyLeuSerlys	169
Db	664	GTTTTACAAACCGGCTGGAAGAATTCGAAGATTTGTCATCAGTACTGCTTCATCTGGCTAAG	723
Qy	170	MetGluGlnAnsglylIleMetSerThrAlacysGlyTrProglyTyrrValAlalaprogu	189
Db	724	CTAGAA---AATGGCCATCATCAGAGCCCTGTGGGACCCCAGAGTACTCTGGCCCCAGAG	780
Qy	190	ValLeuLaGLNLYSPROtyrSerLysAlaValAspcystripSrIIeglyvalIethr	209

```
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 47
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: n=a or g or c or t/u
US-09-935-464-47

Alignment Scores:
Pred. No.: 1-82e-55 Length: 467
Score: 667.00 Matches: 136
Percent Similarity: 93.15% Conservatives: 0
Best Local Similarity: 93.15% Mismatches: 9
Query Match: 26.54% Indels: 1
DB: 9 Gaps: 0

US-09-960-643-2 (1-476) x US-09-935-464-47 (1-467)

Qy 1 MetGlyArgLysGluGluAspCysSerTrpLysGlnThrThrAsnIleArg 20
Db 30 ATGGGTCGAAGGAAGAAGATGACTGCAGTTCTTGGGAAGAAACAGACCACCAACATCCGG 89
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGlyValPheLeuVal 40
Db 90 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCANAAGTTTTTCTCGTG 149
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
Db 150 AAGCAAGACTGACTGGGAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCACCTGCCTTC 209
Qy 61 ArgAspSerLeuGluAsnGluIleAlaValLeuLysLysLysIleLysHisGluAsnIle 80
Db 210 CGGACAGCAGCTGNAGATGAGATGCTGTGTTGAAAAGATCAAGCATGAAAACATP 269
Qy 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
Db 270 GTGACCTGGAGACATCTATGAGAGCACCACCACTACTACTGNTCATGCACTGTT 329
Qy 101 SerGlyGlyLeuPheAsp-ArgIleLeuGluArgGlyValTyrThrGluLysAspAl 120
Db 330 TCTGGTGGGAGCTCTTTGACCGGATCTCTGGAGCGGNGTGTCTACACAGAAGATGC 389
Qy 120 aSerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVa 140
Db 390 CAGNCTGGGTGATCCACANGTCTTTCNCAGTGAATACCTACATGAGATGGCATCGT 449
Qy 140 LHisArgAspLeuLys 145
Db 450 NCACAGAGACTTAAAG 465

RESULT 13
US-09-935-464-48
; Sequence 48, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 48
; LENGTH: 470
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: n=a or g or c or t/u
US-09-935-464-48

Alignment Scores:
Pred. No.: 2-38e-53 Length: 470
Score: 645.00 Matches: 139
Percent Similarity: 93.29% Conservatives: 0
Best Local Similarity: 93.29% Mismatches: 6
Query Match: 25.67% Indels: 4
DB: 9 Gaps: 0

US-09-960-643-2 (1-476) x US-09-935-464-48 (1-470)

Qy 1 MetGlyArgLysGluGluAspCysSerTrpLysLysGlnThrThrAsnIle-Ar 20
Db 30 ATGGGTCGAAGGAAGAAGATGACTGCAGTTCTCTGGAAGAAACAGACCACCAACATCCG 89
Qy 20 gLysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGlu-ValPhe-Leu 39
Db 90 GAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAAAGTTTTCCTG 149
Qy 40 VallysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAla 59
Db 150 GTGAAGCAAAAGACTGACTGGGAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCACCTGCC 209
Qy 60 PheArgAspSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
Db 210 TTCGGGACAGCAGCTGGGAGATGAGATTGCTGTGTTGAAAAGATCAAGCATGAAAAC 269
Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99
Db 270 ATTGTGACCTGGAGGACATCTATGAGAGCACCACCACCTACTACTCTGGTGCATGACGCTT 329
Qy 100 ValSerGlyGlyGluLeuPheAsp-ArgIleLeuGluArgGlyValTyrThrGluLysAs 119
Db 330 GTTCTGTGGGGAGCTCTTTGACCGGATCTCTGGAGCGGNGTGTCTACACAGAAGA 389
Qy 119 pAlaSerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyI 139
Db 390 TGCCAGNCTGGGTGATCCACANGTCTTTCNCAGTGAATACCTACATGAGATGGCAT 449
Qy 139 eValHisArgAspLeuLys 145
Db 450 CGTNCACAGACATTAAAG 468

RESULT 14
US-09-934-406-1
; Sequence 1, Application US/09934406
; Publication No. US20020192204A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 15985, A NOVEL HUMAN SERINE/THREONINE
; FILE REFERENCE: 10448-078001
; CURRENT APPLICATION NUMBER: US/09/934,406
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/226,740
; PRIOR FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3552
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (208)....(2505)
US-09-934-406-1

Alignment Scores:
```



```
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7777
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (10)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (19)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (20)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-7777

Alignment Scores:
Pred. No.:      3 07e-52      Length:      474
Score:          633.50      Matches:      118
Percent Similarity: 87.90%      Conservative: 20
Best Local Similarity: 75.16%      Mismatches:  18
Query Match:      25.21%      Indels:      1
DB:               9          Gaps:      1

US-09-960-643-2 (1-476) x US-09-796-692-7777 (1-474)
QY 115 TyrThrGluLysAspAlaSerLeuValIleGlnValLeuSerAlaValLysTyrLeu 134
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 472 TACACGGAGCGGAGCGCGCGCTCATCTCCAGGTGCTGGATGCTGGAATACCTG 413
QY 135 HisGluAsnGlyIleValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrPro 154
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 412 CATGACCTGGCGATTGTACACCGGGATCTCAAGCCAGAGAACTCTGTGTACTACAGCCTG 353
QY 155 GluGluAsnSerLysIleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly 174
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 352 GATGAGACTCCAAATCAGTCTCCGACTTTGGCCTCTCCAAGATGGAGGCCGGC 293
QY 175 ----IleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGln 193
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 292 AGTGTGCTCTCCACCGCTGTGGAATCCGGGATACGTGGCCCTGAAGTCTTGGCCAG 233
QY 194 LysProTyrSerLysAlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeu 213
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 232 AAGCCCTACAGCAGGCTGTGGATTGCTGGTCCATAGTGTATCGCTACATCTTGCTC 173
QY 214 CysGlyTyrProPheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGlu 233
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 172 TCGGTTACCTCCCTTCTATGACGAGAAATGATGCCAACTCTTTGAACAGATTTCAG 113
QY 234 GlyTyrTyrGluPheGluSerProPheTrpAspPheSerGluSerAlaLysAspPhe 253
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 112 GCCGAGTACGAGTTTGTACTCTCTACTGGGACGACATCTCTGACTCTGCCAAAGATTTC 53
QY 254 IleCysHisLeuLeuGluLysAspProAsnGluArgTyrThrCysGluLys 270
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 52 ATCCGGCACTTGATGGAGAGAGACCCAGAGAAANNNGATTACCNGTGAGCAG 2
```

Search completed: March 15, 2003, 07:54:16  
Job time : 161 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 15, 2003, 05:02:42 ; Search time 64 Seconds  
(without alignments)  
2280.908 Million cell updates/sec

Title: us-09-960-643-2

Perfect score: 2513

Sequence: 1 MGRKEEDCSSKKQTNNR.....VKAGSSHCRAGTGVCLIM 476

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09960643/runat\_07032003\_090517\_19900/app\_query.fasta\_1.647  
-DB=Issued\_Patents\_NA -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LISN=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09960643 -CGN1\_1\_40=runat\_07032003\_090517\_19900 -NCPU=6 -ICPU=3  
-NO\_MLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	990	39.4	1282	2	US-08-878-989-12
2	990	39.4	1282	4	US-09-272-796-12
3	747.5	29.7	3124	4	US-09-734-030-1
4	664.5	26.4	3471	2	US-08-715-568A-2
5	586.5	23.3	1858	4	US-09-529-093A-1
6	578.5	23.0	1735	4	US-09-746-694-3
7	574.5	22.9	2061	4	US-09-800-960-1
8	562.5	22.4	2934	1	US-08-198-446B-16
9	562.5	22.4	2934	2	US-08-870-693-16
10	542.5	21.6	1333	4	US-09-142-551A-1
11	542.5	21.6	2132	2	US-09-159-385-3
12	542.5	21.6	2132	4	US-09-186-277-3

13	537.5	21.4	1400	1	US-08-464-164-1	Sequence 1, Appli
14	537.5	21.4	1400	2	US-08-338-057-1	Sequence 1, Appli
15	537.5	21.4	1400	2	US-08-668-416-1	Sequence 1, Appli
16	531	21.1	2514	3	US-08-655-352-1	Sequence 1, Appli
17	531	21.1	2514	4	US-09-258-016-1	Sequence 1, Appli
18	531	21.1	2514	4	US-09-257-825B-1	Sequence 1, Appli
19	525.5	20.9	4935	2	US-08-631-097-3	Sequence 3, Appli
20	525.5	20.9	5886	4	US-08-810-712-9	Sequence 9, Appli
21	522.5	20.8	1429	2	US-09-159-385-4	Sequence 4, Appli
22	522.5	20.8	1429	4	US-09-186-277-4	Sequence 4, Appli
23	510	20.3	1776	3	US-08-655-352-10	Sequence 10, Appli
24	510	20.3	1776	4	US-09-258-016-10	Sequence 10, Appli
25	510	20.3	1776	4	US-09-257-825B-10	Sequence 10, Appli
26	495.5	19.7	2374	4	US-09-347-801-3	Sequence 3, Appli
27	477	19.0	1929	4	US-09-359-161-4	Sequence 4, Appli
28	460.5	18.3	1742	4	US-08-557-006C-38	Sequence 38, Appli
29	460.5	18.3	2652	4	US-08-557-006C-39	Sequence 39, Appli
30	460.5	18.3	2761	4	US-08-557-006C-24	Sequence 24, Appli
31	459.5	18.3	1736	4	US-08-557-006C-37	Sequence 37, Appli
32	459.5	18.3	1783	4	US-08-557-006C-36	Sequence 36, Appli
33	452	18.0	362	4	US-08-905-223-89	Sequence 89, Appli
34	451.5	18.0	2549	4	US-09-467-082-3	Sequence 3, Appli
35	449.5	17.9	1417	1	US-08-713-828-2	Sequence 2, Appli
36	449.5	17.9	1417	2	US-08-919-627-2	Sequence 2, Appli
37	449.5	17.9	1417	2	US-09-096-245-2	Sequence 2, Appli
38	444	17.7	2637	4	US-09-735-934A-1	Sequence 1, Appli
39	443	17.6	1647	3	US-09-101-146-44	Sequence 44, Appli
40	441.5	17.6	1349	1	US-07-951-715A-20	Sequence 20, Appli
41	441.5	17.6	1349	2	US-08-459-448A-20	Sequence 20, Appli
42	441.5	17.6	1349	3	US-08-459-595A-20	Sequence 20, Appli
43	441.5	17.6	1349	3	US-08-459-504B-20	Sequence 20, Appli
44	441.5	17.6	1349	3	US-08-459-444-20	Sequence 0, Appli
45	441.5	17.6	1349	4	US-09-547-422-20	Sequence 0, Appli

#### ALIGNMENTS

RESULT 1  
US-08-878-989-12  
; Sequence 12, Application US/08878989  
; Patent No. 5885803  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl G.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
; TITLE OF INVENTION: KINASES  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/878,989  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J J

```
;
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
;
;
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT06
; CLONE: 827431
;
US-08-878-989-12

Alignment Scores:
Pred. No.: 1,29e-86 Length: 1282
Score: 990.00 Matches: 193
Percent Similarity: 74.92% Conservative: 49
Best Local Similarity: 59.75% Mismatches: 73
Query Match: 39.40% Indels: 8
DB: 2 Gaps: 2

US-09-960-643-2 (1-476) x US-08-878-989-12 (1-1282)

QY 13 LysLysGlnThrThrAsnIleArgLysThrPheIlePheMetGluValLeuGlySerGly 32
DB 213 AAGAAACACAGCAGGAGCATCAGCAGCGCTCTACGAGATCCCGAGAGCGCTCGGCTCGGGT 272
QY 33 AlaPheSerGluValPheLeuValLysGlnArgLeuThrGlyLysLeuPheAlaLeuLys 52
DB 273 GCCTTCTCGAGGTGCTGCTGCCAGGAGCGGGCTCCGACACCTCGTGGCCCTCAAG 332
QY 53 CysIle---LysLysSerProAlaPheArgAspSerSerLeuGluAsnGluIleAlaVal 71
DB 333 TGCAATCCCAAGAAGCGCTCCGGGGCAAGGAGCGCTGTGGGAGAACGAGATCGCAGTG 392
QY 72 LeuLysLysIleLysGluAsnIleValThrLeuGluAspIleTyrGluSerThrThr 91
DB 393 CTCCTAGATCATCAGTACCCCAACATCGTCGCTGGAGGATGTCACAGAGCCCTTCC 452
QY 92 HisTyrTyrLeuValMetGlnLeuValSerGlyGlyGluLeuPheAspArgIleLeuGlu 111
DB 453 CACCTCTACCTGGCCATGGAATGTGACGGGTGGCGAGCTGTTTGACCGCATCATGGG 512
QY 112 ArgGlyValTyrThrGluLysAspAlaSerLeuValIleGlnValLeuSerAlaVal 131
DB 513 CCGGCTCTCTACAGAGAAGGATGCCAGCATCTGTGGTCAAGTCTTGGCGCCGTC 572
QY 132 LysTyrLeuHisGluAsnGlyIleValHisArgAspLeuLysProGluAsnLeuLysTyr 151
DB 573 TCCTACCTGCACAGCGCTGGGATCGTCACCGGACCTCAAGCCGCAAAACCTCCTGTAT 632
QY 152 LeuThrProGluGluAsnSerLysIleMetIleThrAspPheGlyLeuSerLysMetGlu 171
DB 633 GCCAGCCCTTTGAGGATCGAAGATCATGTCTCTGACCTTGGACTCTCCAAATCCAG 692
QY 172 GlnAsnGlyIleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGluValLeu 191
DB 693 GCTGGAAACATGCTAGGACACCGCTGTGGGACCCCTGGATATGTGGCCCGAGAGCTCTG 752
QY 192 AlaGlnLysProTyrSerLysAlaValAspCysTrpSerIleGlyValIleThrTyrIle 211
DB 753 GAGCAAGAACCTTACGGAAGCCGATAGTGTGGCCCTGGCGGTCATCTCCTACATC 812
QY 212 LeuLeuGlyTyrProPheTyrGluThrGluSerLysLeuPheGluLysIle 231
DB 813 CTGCTGTGGGTACCCCTTCTACGACGAGAGCGACCTGTAGCTCTTCAGCCAGATC 872
QY 232 LysGluGlyTyrTyrGluPheGluSerProPheTrpAspIleSerGluSerAlaLys 251
```

```
DB 873 CTGAGGGCCAGCTATGAGTTGACTNTCTTCTGGGATGACATCTCAGATCAGGCAAA 932
QY 252 AspPheIleCysHisLeuLeuGluLysAspProAsnGluArgTyrThrCysGluLysAla 271
DB 933 GACTTTATTGGCACCTCTTGGAGCGAGACCTTCAGAAGAGGTTCACTGCCAACAGGCC 992
QY 272 LeuSerHisProTrpIleAspGlyAsnThrAlaLeuHisArgAspIleTyrProSerVal 291
DB 993 TTGCGGAGCTTTGGATCTTTTGGACACAGGCTTTGGCAGGACATCTTAGGGTTTGTG 1052
QY 292 SerLeuGlnIleGlnLysAsnPheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAla 311
DB 1053 AGTGAGCAGATCCGGAAGAACTTTGTTGGACACACTGGAAGCGAGCTTCAATGCCACC 1112
QY 312 AlaValValHisHisMetArgLysLeuHisMetAsnLeuHisSerProGlyValArgPro 331
DB 1113 TTGTTCTCGCCACATCCGGAAGCTG-----GGCAGATCCCA 1151
QY 332 GluValGlu 334
DB 1152 GAGGCGGAG 1160
RESULT 2
US-09-272-796-12
; Sequence 12, Application US/09272796
; Patent No. 6207148
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/09/272,796
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/878,989
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT06
; CLONE: 827431
```

```
US-09-272-796-12
Alignment Scores:
Pred. No.: 1,29e-86 Length: 1282
Score: 990.00 Matches: 193
Percent Similarity: 74.92% Conservative: 49
Best Local Similarity: 59.75% Mismatches: 73
Query Match: 39.40% Indels: 8
DB: 4 Gaps: 2

US-09-960-643-2 (1-476) x US-09-272-796-12 (1-1282)
QY 13 LysLysGlnThrThrAsnIleArgLysThrPheIlePheMetGluValLeuGlySerGly 32
||||| ||| :|||
Db 213 AAAAAACACGAGGACATACGACGCGTCTACGAGATCCGCGAGAGCGCTCGGCTCGGT 272
||||| ||| :|||

QY 33 AlaPheSerGluValPheLeuValLysGlnArgLeuThrGlyLysLeuPheAlaLeuLys 52
||||| ||| :|||
Db 273 GCCTTCTCCAGGTGGTGTGCCCCAGGAGCGGGCTCGGCACACCTCGTGGCCCTCAAG 332
||||| ||| :|||

QY 53 CysIle---LysLysSerProAlaPheArgAspSerSerLeuGluAsnGluIleAlaVal 71
||||| ||| :|||
Db 333 TGCATCCCCAAGAAGCCCTCGGGGCAAGGAGGCCCTGGTGAGAACGAGATCGCAGTG 392
||||| ||| :|||

QY 72 LeuLysLysIleLysHisGluAsnIleValThrLeuGluAspIleTyrGluSerThr 91
||||| ||| :|||
Db 393 CTCCTGAGGATCAGTACCCCAACATCGTCTGAGGAGTGTCCAGGAGCGCCTTCC 452
||||| ||| :|||

QY 92 HisTyrTyrLeuValMetGlnLeuValSerGlyGlyGluLeuPheAspArgIleLeuGlu 111
||||| ||| :|||
Db 453 CACCTCTACCTGCCATGGAACCTGTTGACGGTGGCGAGCTGTTGACCGCATCGAG 512
||||| ||| :|||

QY 112 ArgGlyValTyrThrGluLysAspAlaSerLeuValIleGlnValLeuSerAlaVal 131
||||| ||| :|||
Db 513 CGGGGCTCTACACAGAGAGATGCCACCATCTGGTGGGTGACAGTCTTGGCGCGTC 572
||||| ||| :|||

QY 132 LysTyrLeuHisGluAsnGlyIleValHisArgAspLeuLysProGluAsnLeuTyr 151
||||| ||| :|||
Db 573 TCCTACCTGCACAGCTGGGATCGTGCACGGGACCTCAGCCCGAANAACCTCTGTAT 632
||||| ||| :|||

QY 152 LeuThrProGluGluAsnSerLysIleMetIleThrAspPheGlyLeuSerLysMetGlu 171
||||| ||| :|||
Db 633 GCACACCCCTTTGAGGACTCGAAGATCATGCTCTGACTTTGGACTCTCCAAAATCCAG 692
||||| ||| :|||

QY 172 GlnAsnGlyIleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGluVal 191
||||| ||| :|||
Db 693 GCTGGGAACATGCTAGGACCCCTGCTGGACCCCTGTGATATGTGGCCCGAGAGCTGTG 752
||||| ||| :|||

QY 192 AlaGlnLysProTyrSerLysAlaValAspCysTyrPheIleGlyValIleThrTyrIle 211
||||| ||| :|||
Db 753 GAGCAGAAACCTACGGGAAGCCGTAGATGTGTGGGCCCTGGGGGTCTATCTCTACATC 812
||||| ||| :|||

QY 212 LeuLeuCysGlyTyrProPheThrGluGluThrGluSerLysLeuPheGluLysIle 231
||||| ||| :|||
Db 813 CTGCTGTGGGTACCCCTTCTACGACGAGACGACCCCTGAGCTCTTCAGCCAGATC 872
||||| ||| :|||

QY 232 LysGluGlyTyrTyrGluPheGluSerProPheThrAspAspIleSerGluSerAlaLys 251
||||| ||| :|||
Db 873 CTCAGGGCCAGCATGAGTTGACTTCTCTTCTGGGTGATGACATCTCAGATCAGGCA 932
||||| ||| :|||

QY 252 AspPheIleCysHisLeuLeuGluLysAspProAsnGluArgTyrThrCysGluLysAla 271
||||| ||| :|||
Db 933 GACTTTATTCGGCAGCTCTGAGGAGACCTTCAGAGAGGTTTCACTGACCAACAGGCC 992
||||| ||| :|||

QY 272 LeuSerHisProTrpIleAspGlyAsnThrAlaLeuHisArgAspIleTyrProSerVal 291
||||| ||| :|||
Db 993 TTGCGGAGCTTTGGATCTTTGGACACACAGGCTTTGCGACGGGACATCTTAGGTTTGT 1052
||||| ||| :|||

QY 292 SerLeuGlnIleGlnLysAsnPheAlaLysSerLysTyrPheGlnAlaPheAsnAlaLa 311
||||| ||| :|||
Db 1053 AGTGACGACATCGGAGAGACTTGTCTGGACACACTCGAGCGAGCCTCTCAATGCCACC 1112
||||| ||| :|||

QY 312 AlaValValHisMetArgLysLeuHisMetAsnLeuHisSerProGlyValArgPro 331
||||| ||| :|||
```



```
|||||...
818 CTGGCCAGTGTCCGACAGAGGCTGACCTGTGTGAGAACCCACCTGTGGCAGCCT 877
QY GlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaValAspCysTrp 203
Db GAGTACATTGCCAGAGTCTGTGTCGCAAGCCATACCAACTCAGTGGACATGTGG 937
QY SerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyrGluGluThr 223
Db ::::::||||| 1111111111 111 111111 ::::::
938 CGCTGGCGGTCATGCTCATCTACTCAGTGGCACCACCTGCTTGGAGGATGACAAC 997
QY 224 GluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSerProPheTrp 243
Db ::::::||||| 1111111111 111 111 111 :::
998 CGTACCGGCTGTACCGGCAGATCTCAGGGCAAGTACAGTACTCTGGGGAGCCCTGG 1057
QY 244 AspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuLysAspProAsn 263
Db ::::::||||| 1111111111 111111 ::::::
1058 CCTAGTGTGTCAACCTCGGCAAGGACTTCATTGACCGCTGTGACAGTGGACCCCTGGA 1117
QY 264 GluArgTyrThrCysGluLysAlaLeuSerHisProThrIleAspGlyAsnThrAlaLeu 283
Db ||| ||| ::::::||||| 1111111111 ::::::
1118 GCCGTGATGACTGACGTGCGAGCCCTGAGGCACCCGGGTG----- 1159
QY 284 HisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAlaLysSerLys 303
Db ||| ||| ::::::||||| 111111 ::::::
1160 -----GTGAGCATG----- 1168
QY 304 TrpArgGlnAlaPheAsnAlaAlaValHisMetArgLysLeuHisMetAsn 323
Db ||| ||| ::::::||||| 111111 :::
1169 -----GTCGCTCTTCATCCATGAAGAACCCTGCACCGCTCC 1204
QY 324 LeuHisSerProGlyValAlaProGluValGluAsnArgProProGluThrGlnAlaSer 343
Db ::: ||| ::::::||||| 111111 ::::::
1205 ATA---TCCAGAACCTCTTAACGTGCTCTCGGCTGCCAGAGCCACCAATCTGCC 1261
QY 344 GluThrSerArg-ProSerSerProGluIleThrIleThrGluAlaProValLeuAspH1 363
Db ::::::||||| 1111111111 111111 ::::::
1262 CAGTCCACGGCTCCAGCCGCTCCA-----CAGCCTCCAATA---AGTCA 1303
QY 363 sSerVal-AlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArgProThra 383
Db ||| ||| ||| ::::::||||| 111 ::::::
1304 CGCCGTGCGGGGAACGGAGCTCG-----GGAGCTCAACCTCGCG 1345
QY 383 laProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSerSerL 403
Db ||| ::::::||||| 111 ::::::
1346 TACCAGCA-----GCAATCAATGGCTGAGCCGATGGCTGTCACACATGCAGCACGA 1399
QY 403 euValProMethHis-----GlnGlySerLeuAlaAlaGlyP 415
Db ||| ||| ||| ::::::||||| 111111 ::::::
1400 CCCAGCCTGGCCACACACTGTGTGTCATCTGGGTCCGATGCCCTCTCTGGAGATAGGCG 1459
QY 415 roCysGly 417
Db ||| ||| ||| ::::::||||| 111111 ::::::
1460 TATGTGGC 1467

RESULT 5
US-09-529-093A-1
; Sequence 1, Application US/09529093A
; Patent No. 6413755
; GENERAL INFORMATION:
; APPLICANT: LUYTEN, Walter H.M.L.
; APPLICANT: PARKER, Andrew E.
; APPLICANT: MCGOWAN, Clare H.
; APPLICANT: BLASINA, Alessandra
; TITLE OF INVENTION: Human Checkpoint Kinase, hCDS1, Compositions and Methods
; FILE REFERENCE: TSRI 649.0
; CURRENT APPLICATION NUMBER: US/09/529,093A
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: PCT/EP98/06981
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: GB 9722320.0
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 6
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1858
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (66)..(1694)
US-09-529-093A-1

Alignment Scores:
Pred. No.: 3,17e-47 Length: 1858
Score: 586.50 Matches: 142
Percent Similarity: 52.60% Conservative: 60
Best local Similarity: 36.98% Mismatches: 131
Query Match: 23.34% Indels: 51
DB: 4 Gaps: 11

US-09-960-643-2 (1-476) x US-09-529-093A-1 (1-1858)
QY 7 AspAspCysSerSerTrpLysLysGlnThrThrAsnIleArgLysThrPheIlePheMet 26
Db 684 GATGATCAGTCAGTTTATCTTAAGGCA-----TTAAGAGATGAATACATCATGTCA 734
QY 27 GluValLeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGly 46
Db 735 AAAAATCTTGGAAAGTGGTGGCTGTGGAGAGGTAAAGCTGGCTTCGAGAGGAACACATGT 794
QY 47 LysLeuPheAlaLeuLysCysIleLysLys----- 56
Db 795 AAGAAGTAGCCATAAAGATCATCAGCAAAAGGAAGTTTGCCTATTTGGTTTCAGCAAGAGAG 854
QY 57 ---SerProAlaPheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIle 75
Db 855 CGAGACCAGCTCTC-----AATGTTGAAACAGAAATAGAAATTTTGAAGAAAGCTA 905
QY 76 LysHisGluAsnIleValThrLeuGluAspIleTyrGluSerThrThrHisTyrLeu 95
Db 906 AATCATCTTCATCATCAAGATTAACACTTTTGTGATGCA---GAAGATTATTATATT 962
QY 96 ValMetGlnLeuValSerGlyGlyLeuPheAspArgIleLeuGluArgGlyValTyr 115
Db 963 GTTTTGGAAATTGATGAAAGGGGAGAGCTGTTTGAACAAAGTGGTGGGAATAAACGCCGTG 1022
QY 116 ThrGluLysAspAlaSerLeuValIleGlnValLeuSerAlaValLysTyrLeuHis 135
Db 1023 AAAGAAGTACCTCGCAAGCTCTATTTTACCAGATGCTCTTGGCTGTGCAGTACCTTCAT 1082
QY 136 GluAsnGlyIleValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGlu 155
Db 1083 GAAACGGTATTATACACCGTGACTTAAGCCAGAGAAATGTTTACTGTCAATCAAGAA 1142
QY 156 GluAsnSerLysIleMetIleThrAspPheGlyLeuSerLys---MetGluGlnAsnGly 174
Db 1143 GAGGACTGCTTATAAAGATTACTGATTTTGGGCACCTCCAAGATTTTGGGAGAGACCTCT 1202
QY 175 IleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGln--- 193
Db 1203 CTCATGAGAACCCTTATGTGAACCCCCACCTTGGCGCTGAGGTTTGTGTTCTGCTT 1262
QY 194 -----LysProTyrSerLysAlaValAspCysTrpSerIleGlyValIleThrTyrIle 211
Db 1263 GGGACTGTGGGTATAACCGTGCTGTGCAGCTGTGGAGTTTAGGAGTTATTCTTTTATC 1322
QY 212 LeuLeuCysGlyTyrProPheTyrGlu---GluThrGluSerLysLeuPheGluLys 230
Db 1323 TGCCTTAGTGGGTATCCACCTTCTCTGAGCATAGGACTCAAGTGTCACTGAAGGATCAG 1382
QY 231 IleLysGluGlyTyrTyrGluPheGluSerProPheTrpAspIleSerGluSerAla 250
Db 1383 ATCCACCATGGAAAATCAACACTTCATTCCTCCTGAGTGTGGCGAAGTCTCAGAGAAAGCT 1442
QY 251 LysAspPheIleCysHisLeuLeuLysAspProAsnGluArgTyrThrCysGluLys 270
```

```

Db 1443 CTGGACCTTGTCAAGAAGTTGTTGGTAGTCACCAAGCCACTTTTACGACAGAGAA 1502
      |||   ...   |||||   |||||   |||...|||   |||...
Qy 271 AlaLeuSerHisProtrpIleAspGlyAsnThrAlaLeuHisArgAspIleYrProSer 290
      |||||   |||||   |||||   |||||   |||||   |||||   |||||   |||||
Db 1503 GCCTTAAGACACCGTGG-----1520
Qy 291 ValSerLeuGlnIleGlnLysAsnPheAlaLysSerLysTrpArgGlnAlaPheAsnAla 310
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 1521 -----CCTCAGGATGAGACATGAGAGAAAGTTTCAAGATCTTCTGTCTGAG 1568
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 311 AlaAlaValValHisMetArgLysLeuHisMetAsnLysSerProGlyVal--- 329
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 1569 GAA-----AATGAATCCACAGCTCTACCCAGGTTCTAGCCCGCTTCTACTACT 1619
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 330 -----ArgProGluValGluAsnArgProProGluThrGlnAlaSerGluThrSerArg 347
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 1620 CGAAAGCGGCC-----CGTGAAGGGGNAAGCCGAGGGTCCCGAGACCACAAAG 1667
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 348 ProSerSerProGluIleThrIleThrGluAlaProValLeuAspHisSerValAlaLeu 367
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 1668 CGCCCACTGTGCTGCTGCTGTGTTCTGAACCTCCGTGTTGAACACCAAGAAATGTA 1727
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 368 ProAlaLeuThr 371
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 1728 CCTTCTTTCACT 1739

RESULT 6
US-09-746-694-3
; Sequence 3, Application US/09746694
; Patent No. 6451538
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF CHK2 EXPRESSION
; FILE REFERENCE: RTS-0228
; CURRENT APPLICATION NUMBER: US/09/746,694
; CURRENT FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO. 3
; LENGTH: 1735
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1632)
US-09-746-694-3

Alignment Scores:
Pred. No.: 1.69e-46 Length: 1735
Score: 578.50 Matches: 126
Percent Similarity: 60.28% Conservative: 47
Best Local Similarity: 43.90% Mismatches: 91
Query Match: 23.02% Indels: 23
DB: 4 Gaps: 7

US-09-960-643-2 (1-476) x US-09-746-694-3 (1-1735)

Qy 7 AspAspCysSerSerTrpLysLysGlnThrThrAsnIleArgLysThrPhePheMet 26
      |||||   |||   :|||   :|||   :|||   :|||   :|||   :|||
Db 619 GATGATCAGTCAGTTTATCTCAAGGCA-----TTAAGAGATGAATACATCATGTCA 669
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 27 GluValLeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGly 46
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 670 AAAACTCTTGAAGTGTGCTGTGGAGAGTTAAAGCTGGCTTCGAGAGAGAAACATGT 729
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 47 LysLeuPheAlaLeuLysCysIleLysLys-----56
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 730 AAGAAAGTAGCCATAAAGATCATCAGCAAAAGGAAGTTTGCTATTGGTTCAGCAAGAGAG 789
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 57 ---SerProAlaPheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIle 75
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 790 GCAGACCCAGCTCTC-----AATGTTGAACACAGAAATAGAAATTTTGAAAGAGCTA 840
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
```

```

Qy 76 LysHisGluAsnIleValThrLeuGluAspIleYrGluSerThrThrHisYrYrLeu 95
      |||   |||   :|||   :|||   :|||   :|||   :|||   :|||
Db 841 AATCATCTTGCATCATCAAGATTAAACACTTTTGTATGCA---GAAGATTATTATATT 897
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 96 ValMetGlnLeuValSerGlyGlyLeuPheAspArgIleLeuGluArgGlyValYr 115
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 898 GTTTTGAATTGATGAAGGGGAGAGCTGTTTCACAAAGTGGTGGGAATAAACGCCTG 957
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 116 ThrGluLysAspAlaSerLeuValIleGlnValLeuSerAlaValLysYrLeuHis 135
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 958 AAAGAAGTACTCTCAAGCTCTATTTTACCAGATGCTGTGGCTGTGCATCTCAT 1017
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 136 GluAsnGlyIleValHisArgAspLeuLysProGluAsnLeuLeuYrLeuProGlu 155
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 1018 GAAACGGTATTATACACCGTGACTTAAAGCCAGAGAAATGTTTACTGTCTCAAGAA 1077
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 156 GluAsnSerLysIleMetIleThrAspPheGlyLeuSerLys---MetGluGlnAsnGly 174
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 1078 GAGGACTCTCTATAAAGATTACTGATTTGGGCACTCCAAGATTTTGGGAGAGACCTCT 1137
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 175 IleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGln--- 193
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 1138 CTCATGACAACCTATATGGAACCCCACTACTTGGCGCTGAAGTCTGTGTTCTGTT 1197
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 194 -----LysProTyrSerLysAlaValAspCysTrpSerIleGlyValIleThrYrIle 211
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 1198 GGGACTGCTGGGTATAACCGTCTGTGGACTGCTGGAGTTTAGGAGTATTCTTTTATC 1257
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 212 LeuLeuCysGlyYrProProPheYrGlu---GluThrGluSerLysLeuPheGluLys 230
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 1258 TGCCTTAGTGGGTATCCACCTTCTCTGAGCATAGGACTCAAGTGTCTCAAGAGATCAG 1317
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 231 IleLysGluGlyYrYrGluPheGluSerProPheTrpAspAspIleSerGluSerAla 250
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 1318 ATCACCAGTGGAAATACAACTTCATCTGAAGTCTGGGCAAGTCTCAGAGAAAGCT 1377
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 251 LysAspPheIleCysHisLeuLeuLysAspProAsnGluArgYrThrCysGluLys 270
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 1378 CTGGACCTTGTCAAGAAGTTTGTGTAGTGGATCCAAAGGCACGTTTACGACAGAA 1437
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Qy 271 AlaLeuSerHisProTrpIle 277
      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
Db 1438 GCCTTAAGACACCCGTGGCTT 1458

RESULT 7
US-09-800-960-1
; Sequence 1, Application US/09800960
; Patent No. 6387677
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001158
; CURRENT APPLICATION NUMBER: US/09/800,960
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2061
; TYPE: DNA
; ORGANISM: Human
US-09-800-960-1

Alignment Scores:
Pred. No.: 5.44e-46 Length: 2061
Score: 574.50 Matches: 149
Percent Similarity: 48.89% Conservative: 72
Best Local Similarity: 32.96% Mismatches: 193
Query Match: 22.86% Indels: 40
DB: 4 Gaps: 9

US-09-960-643-2 (1-476) x US-09-800-960-1 (1-2061)
```

QY 17 ThrAsnIleArgLysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGlu 36  
DB 134 ACCGGTTTACCGCAGCAGTACAGCTCTTCGAGGAGCTTGGCAAGGGTCTTCTCTGTG 193  
QY 37 ValPheLeuValLysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle----- 54  
DB 194 GTCGCGAGGTGTGTGAAGAAACCTCCACGCGAGGAGTAGCGCAAAATCATCAATACC 253  
QY 55 LysLysSerProAlaPheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLys 74  
DB 254 AAGAAATTGCTGCCCGGGATCACAGAACTAGACGTGAGGCTCGGNATGTCGACTT 313  
QY 75 IleLysHisGlnLeuValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyr 94  
DB 314 CTGAACATCAAAACATCGTGGCCCTCCATGACAGTATTCTGAAGAGGGTTTCACTAC 373  
QY 95 LeuValMetGlnLeuValSerGlyGlyLeuLeuPheAspArgIleLeuGluArgGlyVal 114  
DB 374 CTCGTGTTTGACCTGTGTACCGGGGGAGCTGTTGAAGACATGTGTGCCAGAGAGTAC 433  
QY 115 TyrThrLysAspAlaSerLeuValIleGlnValLeuSerAlaValLysTyrLeu 134  
DB 434 TACAGTGAACAGATGCCAGCCACTGTATACATCAGATTCTGGAGAGTGTAAACCAATC 493  
QY 135 HisGluAsnGlyIleValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuPro 154  
DB 494 CACCAGCATGACATCGTCCACAGGACCTGAAGCTTGAGAACCTGCTGTGGCGAGTAA 553  
QY 155 GluLysAsnSerLysIleMetIleThrAspPheGlyLeuSer-----LysMetGlu 171  
DB 554 TGCAGGGTCCCGCGTGAAGCTGGCTGATTGCGCTAGCCATCGAAGTACAGGAGAG 613  
QY 172 GlnAsnGlyIleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGluValLeu 191  
DB 614 CAGCAGCTGTTGGTGTGTTTGTCT--GGCACCCAGGTTACTGTGCCCTGAGGTCTTG 670  
QY 192 AlaGlnLysProTyrSerLysAlaValAspCysTyrSerIleGlyValIleThrTyrIle 211  
DB 671 AGAAAGATCCCTATGAAACCTGTGGATATCTGGGCTGGGGTCTATCTGTATATATC 730  
QY 212 LeuLeuCysGlyTyrProProPheTyrGluThrGluSerLysLeuPheGluLysIle 231  
DB 731 CTCCTGTGGGCTATCTCTCTCTGGATGAGGATCAGCACAAAGCTGTATCAGCAGATC 790  
QY 232 LysGluGlyTyrTyrGluPheGluSerProPheTyrPaspAspIleSerGluSerAlaLys 251  
DB 791 AAGCTGGAGCTATGATTTCCATCACCAGATGGGACACGGTAACCTCTCTGAAGCCAA 850  
QY 252 AspPheIleCysHisLeuLeuGluLysAspProAsnGluArgTyrThrCysGluLysAla 271  
DB 851 AACTGTATCAACAGATCTGACCAATAAACCCAGCAAGCGATCAGGCTGACCAAGGT 910  
QY 272 LeuSerHisProTyrPaspGlyAsnThrAlaLeuHisArgAspIleTyrProSerVal 291  
DB 911 CTCAAGCACCGGTGGTCTGCAACGATCCACGCTGGCATCCATGATGCTGCTCAGGAG 970  
QY 292 SerLeuGlnIleGlnLysAsnPhe--AlaLysSerLysTyrArgGlnAlaPheAsnAla 310  
DB 971 ACTGTGGAGGTGTTGGCAAGTCAATGCCCCGGAGAAACTGAAGGGTCCCATCTCTCAG 1030  
QY 311 AlaAlaValValHisMetArgLysLeuHisMetAsnLeuHisSerProGlyValArg 330  
DB 1031 ACCATGCTGTCTCCAGAACTTCTCAGTTGGCAGGACGAGCTCCGCCCC--GCCTCG 1087  
QY 331 ProGluValGluAsnArgProProGluThrGlnAlaSerGluThrSer----- 346  
DB 1088 CTGCGCGC-GAGCGCCCGCGCTGCGCGGCAAGCTGCGCAAAAGCCTATTGAACAAGAA 1146  
QY 347 -----ArgProSerSerProGluIleThrIleThrGluAlaProValLeuAspHis 363  
DB 1147 GTCGATGGCGGTGTCAGAAAGGAAGTTCAGTTCCAGGTCACCTAA-TGGAGCCAC 1205

QY 364 SerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArgProThrAla 383  
DB 1206 AAACCACTGTGGTACACACAGCTACAGATGGGATCAAGGGCTCCACAGAGAGCTGCAACA 1265  
QY 384 ProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSerSerLeu 403  
DB 1266 CCACCAAGAG----- 1277  
QY 404 ValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSerSerCysLeu 423  
DB 1278 -----ATGAGGACCTCAAGCTGCCGCTCGGCATGGGAATGGCAGCTCGGTGCTG 1331  
QY 424 AsnIleGlySerLysGlyLysSer-----SerTyrCysSerGluProThrLeu 439  
DB 1332 AAGCAGGAGCTCCGGGACAGAACAGCCCCCTCTCAGGCATGCGAGC---CCAGCCTT 1388  
QY 440 LeuLysLysAlaAsnLysLysGlnAsnPheLysSer 451  
DB 1389 CTCTCTCTCTCAGCCATCGGAACAGGAGATCA 1424  
RESULT 8  
US-08-198-446B-16  
; Sequence 16, Application US/08198446B  
; Patent No. 5674996  
; GENERAL INFORMATION:  
; APPLICANT: Hartwell, Leland H.  
; APPLICANT: Weinert, Ted A.  
; APPLICANT: Pilon, Sharon E.  
; APPLICANT: Groudine, Mark T.  
; TITLE OF INVENTION: Cell Cycle Checkpoint Genes  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC  
; STREET: 1420 Fifth Ave., Suite 2800  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101-2347  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/198,446B  
; FILING DATE: 18-FEB-1994  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Shelness, Diana K.  
; REGISTRATION NUMBER: 35,356  
; REFERENCE/DOCKET NUMBER: FHCRI7537  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-682-8100  
; TELEFAX: 206-224-0779  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2934 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; DESCRIPTION: yeast MEC2 cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Saccharomyces cerevisiae  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 395..2724  
US-08-198-446B-16  
Alignment Scores: 1.39e-44 Length: 2934  
Pred. No.:

Score: 562.50 Matches: 157  
Percent Similarity: 48.50% Conservative: 70  
Best Local Similarity: 33.55% Mismatches: 165  
Query Match: 22.38% Indels: 78  
DB: 1 Gaps: 15

US-09-960-643-2 (1-476) x US-08-198-446B-16 (1-2934)

```
QY 10 SerSerTriPlysLysGlnThrThrAsnIleArgLysThrPhe---IlePheMetGluVal 28
Db 944 TCATCAATGGTGCACCAAGACTGCTATTTTAAGGATTTTCGATATTATTCACCAAGTG 1003
QY 29 LeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnAlaThrGlyLysLeu 48
Db 1004 GTGGCCAGGGTGCAATTCGCCAGTAAAGAAAGCCATTGAAAGAACTACTGGGAAACA 1063
QY 49 PheAlaLeuLysCysIleLysLysSerProAlaPheArgAsp---SerSerLeuGluAsn 67
Db 1064 TTCGCGGTGAAGATTATAGTAACCCAAAGTAATAGGCAATATGGATGGTGTGACAGA 1123
QY 68 GluIleAlaValLeuLysLysIleLysHisGluAsnIleValThrLeuGluAspIleTyr 87
Db 1124 GAGTTAGAGTATTGCAAAAGCTCAATCATCAAGGATAGTACGATTGAAAGGATTTTAT 1183
QY 88 GluSerThrThrHisTyrTyrLeuValMetGlnLeuValSerGlyClyLeuPheAsp 107
Db 1184 GAAGTACTGAGAGTATTATATGTTGATGGTGGTCTGTTCTGGTGGTCAATATGGAT 1243
QY 108 ArgIleLeuGluArgGlyValTyrThrGluLysAspAlaSerLeuValIleGlnVal 127
Db 1244 TTGTTGCTGCTCATGGTGGTGGAGAGATGCTGGGAGGAGATATCCAGGAGATA 1303
QY 128 LeuSerAlaValTyrLeuHisGluAsnGlyIleValHisArgAspLeuLysProGlu 147
Db 1304 CTCACAGCAATAAATACATTCATCTATGGGCATCAGCCATCGTACCTAAAGCCGAT 1363
QY 148 AsnLeuLeuTyrLeuThrProGluGluAsnSerLysIleMetIleThrAspPheGlyLeu 167
Db 1364 AATATCTCT---ATTGAACAAGCATCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1420
QY 168 SerLysMetGluGlnAsnGly---IleMetSerThrAlaCysGlyThrProGlyTyrVal 186
Db 1421 GCAAAAGTACAAGAAATGGTCTTTTATGAAACCTTCTGTGGCACITTTGTCATATGTG 1480
QY 187 AlaProGluValLeuAlaGlnLys----- 194
Db 1481 GCACCTGAAGTCAATCAGAGTAAGATACATCCGTATCTCTGATCAATACGAAAGAG 1540
QY 195 ---ProTyrSerLysAlaValLysCysTyrSerIleGlyValIleThrTyrIleLeuLeu 213
Db 1541 AATGAGTACTCTTCGTTAGTGGATATGCTGTAATGGGATGCTTGTGTATGTTATCCTA 1600
QY 214 CysGlyTyrProPheTyrGluThrGluSerLysLeuPheGluLysIleLysGlu 233
Db 1601 ACGGCCACTTACCTTTTAGTGTACACACAGAGGACCATATATAACAGATTGGAAGA 1660
QY 234 GlyTyrThrGluPheGluSerProPheTyrAsp---AspIleSerGluSerAlaLysAsp 252
Db 1661 GGCTCATATCAT---GAAGGGCCCTCAAGATTTCCGGATATCTCGAAGCAAGAGAT 1717
QY 253 PheIleCysHisLeuGluLysAspProAsnGluArgTyrThrCysGluLysAlaLeu 272
Db 1718 TCCATAGATTGTTTACAGGGTGGATCCAAATATAGATTCGACAGCTGCAAAAGCCCTG 1777
QY 273 SerHisProTyrIleAspGlyAsnThrAla-LeuHisArg----- 285
Db 1778 AATCATCCCTGGATCAAGATGATGCTTGGCTCACATCATATATGGTATTTTCACAA 1837
QY 286 -----AspIleTyrProSerValse 292
Db 1838 ATATCCTTTATCAATCGTTGTGCGACAGAGAAATTTATGAAATATGACGATGCTCAA 1897
QY 292 rLeuGlnIleGlnLysAsnPheAlaLysSerLysIleTyrArgGlnAlaPheAsnAlaAla 312
```

```
Db 1898 TACGAATTTGTCAAAGCGCAAGAAATTTACAATATGGACCAACAACCTTCAAG----- 1949
QY 312 aValValHisHisMetArgLysLeuHisMet----- 322
Db 1950 -----AACAGGATCAGGAAGACCAAGATGGAAATTTCAAGATTTAAAAATACCGCA 2002
QY 323 -----AsnLeuHisSerPro---GlyValArgProGluValGluAsnArg--- 336
Db 2003 CACGCCCTATTTCATATACACACGCCCAAAAGCATTTGAAGCAGAACTAGAGAACAAAA 2062
QY 337 -----ProProGluThrGlnAlaSerGluThrSerArgProSerSerProGluL 353
Db 2063 CTTTTCATCTCCAT- AATACTGAGATTCGAAGAGCTCAAGAGCTCAAGAAAGGTTAATGGTAG 2121
QY 353 eThrIleThrGluAlaProValLeuAspHisSerValLeuProAlaLeuThrGlnLe 373
Db 2122 GTTTTACCTTAAACCATTCCTGACAGCATTTATTCAA-----GAAAGCCT 2169
QY 373 uProCysGlnHisGlyArgArgProThrAlaProGly---GlyArgSerLeuAsnCysLe 392
Db 2170 GGAGATTCAGCAAGGTGTGAATCCATTTTTCATTTGTTAGATCCGAGGATTCATTTGTA 2229
QY 392 uValAsnGlySerLeuHisIleSerSerSerSerLeuValProMetHisGlnGlySerLeuAl 412
Db 2230 AATTGAAGACAATAGGTT-GTCTCGAGTTCATTCCTTCATTTTCAAAAAGAGGCA----- 2283
QY 412 aAlaGlyProCysGlyCysCys 419
Db 2284 -----TGCTGT 2289

RESULT 9
US-08-870-693-16
; Sequence 16, Application US/08870693
; Patent No. 5866338
; GENERAL INFORMATION:
; APPLICANT: Hartwell, Leland H.
; APPLICANT: Weinert, Ted A.
; APPLICANT: Pilon, Sharon E.
; APPLICANT: Groudine, Mark T.
; TITLE OF INVENTION: Cell Cycle Checkpoint Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/870,693
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/198,446
; FILING DATE: February 18, 1994
; APPLICATION NUMBER: PCT/US93/04458
; FILING DATE: May 12, 1993
; APPLICATION NUMBER: US 07/884,426
; FILING DATE: May 14, 1992
; APPLICATION NUMBER: US 07/882,051
; FILING DATE: May 12, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelness, Diana K.
; REGISTRATION NUMBER: 35,356
; REFERENCE/DOCKET NUMBER: FPCR110798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-682-8100
```

TELEFAX: 206-224-0779					
INFORMATION FOR SEQ ID NO: 16:					
SEQUENCE CHARACTERISTICS:					
LENGTH: 2934 base pairs					
TYPE: nucleic acid					
STRANDEDNESS: single					
TOPOLOGY: linear					
MOLECULE TYPE: cDNA to mRNA					
DESCRIPTION: yeast MEC2 cDNA					
HYPOTHETICAL: NO					
ANTI-SENSE: NO					
ORIGINAL SOURCE:					
ORGANISM: Saccharomyces cerevisiae					
FEATURE:					
NAME/KEY: CDS					
LOCATION: 395..2724					
US-08-870-693-16					
Alignment Scores:					
Pred. No.:	1,39e-44	Length:	2934		
Score:	562.50	Matches:	157		
Percent Similarity:	48.50%	Conservative:	70		
Best Local Similarity:	33.55%	Mismatches:	165		
Query Match:	22.38%	Indels:	78		
DB:	2	Gaps:	15		
US-09-960-643-2 (1-476) x US-08-870-693-16 (1-2934)					
Qy	10	SerSerTrpLysGlnThrAsnIleArgLysThrPhe---	IlePheMetGluVal	28	
Db	944	TCATCAATGGTGGCCAACAAAGACTGTGATTTTTAAGGATTTTTCGATTATTACGAAGTG	1003		
Qy	29	LeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGlyLysLeu	48		
Db	1004	GTGGCGCAGGGTGCATTTGCCACAGTAAGAAGCCATTGAAGAAGACTACTCTGGGAAAACA	1063		
Qy	49	PheAlaLeuLysCysIleLysLysSerProAlaPheArgASP---	SerSerLeuGluAsn	67	
Db	1064	TTCGCGTGAAGATTATAAGTAAACCAACCAAGTAATAGGCAATATGGATGGTGTGCAACAAG	1123		
Qy	68	GluIleAlaValLeuLysLysIleLysHisGluAsnIleValThrLeuCluaSpIleYr	87		
Db	1124	GAGTTAGAGATTGCAAGAGCTCAATCATCCAAGTAGTAGGATTGAAGGATTTTAT	1183		
Qy	88	GluSerThrThrHisTyrrLeuValMetGlnLeuValSerGlyGlyGluLeuPheAasp	107		
Db	1184	GAAGATACAGAGATTATTATATATGGTGATGGAGTTCGTTCTGGTGGTACCTTAATGGAT	1243		
Qy	108	ArgIleLeuGluArgGlyValTyrrThrGluLysAspAlaSerLeuValIleGlnVal	127		
Db	1244	TTTGTGTGCTGCTCATGGTGGTGGAGAAAGATGCTGGGAGGAGATATCCAGGCAGATA	1303		
Qy	128	LeuSerAlaValLysTyrrLeuHisGluAsnGlyIleValHisArgAspLeuLysProGlu	147		
Db	1304	CTCACAGCAATAAATACATTCACTATATGGGCATCAGCCATCGTACCTAAGCCCGAT	1363		
Qy	148	AsnLeuLeuTyrrLeuThrProGluGluAsnSerLysIleMetIleThrAspPheGlyLeu	167		
Db	1364	AATATTCTT---ATTGAACAAGACGATCTGTATTGTTGAAGATACCACACTTTGGTCTG	1420		
Qy	168	SerLysMetGluGlnAsnGly----IleMetSerThrAlaCysGlyThrProGlyTyrrVal	186		
Db	1421	GCAAAAGTAGAAGAAATGGTCTTTTATGAANAACCTCTCTGGGCACCTTGGCATATGTG	1480		
Qy	187	AlaProGluValLeuAlaGlnlys-----	194		
Db	1481	GCACCTGAAGTCATCAGAGTAAGAATACATCCGTATCTCCTGATGAATACGAAGAAGG	1540		
Qy	195	--ProTyrrSerLysAlaValAaspCysTrpSerIleGlyValIleThrTyrrIleLeu	213		
Db	1541	AARTGAGTACTCTTCGTTAGTGATATGTCTCAATGGGATGTCTGTGATTGTATPCCTPA	1600		
Qy	214	CysGlyTyrrProProPheTyrrGluThruThrGluSerLysLeuPheGluLysIleLysGlu	233		

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-142-551A-1

Alignment Scores:
Pred. No.: 3,46e-43 Length: 1333
Score: 542.50 Matches: 135
Percent Similarity: 52.02% Conservative: 71
Best Local Similarity: 34.09% Mismatches: 139
Query Match: 21.59% Indels: 51
DB: 4 Gaps: 10

US-09-960-643-2 (1-476) x US-09-142-551A-1 (1-1333)
Qy 27 GluValLeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGly 46
Db 179 CAGGTGCTGGCGCTGGGTGTGAACGCAAGAGTGTGGAGTGTTCCTCCTGCGGCGACTGA 238
Qy 47 LysLeuPheAlaLeuLysCysIleLysLysSerProAlaPheArgAspSerSerLeuGlu 66
Db 239 CAGAAGTGTGCCCTGAAGCTCCTGTATGACAGCCCAAGCCCGG-----283
Qy 67 AsnGluIleAlaValLeuLysLysIleLysHisGlu-----Asn 79
Db 284 -----CAGGAGGTAGACCATCATCTGCAGGCTTCTGGCGGCCCCCAT 325
Qy 80 IleValThrLeuGluAspIleTyrGluSerThrHis-----TyrTyrLeu 95
Db 326 ATTGCTGTGATCTGATGTGTATGAGAACATGCACCATGGCAAGCGTGTCTCTCATC 385
Qy 96 ValMetGlnLeuValSerGlyGlyGluLeuPheAspArgIleLeuGly-----113
Db 386 ATCATGAATGCATGCAAGGTGTGTGTTTCAGCAGGATTTCAGGAGCGTGGCGACCA 445
Qy 114 ValTyrThrLysAspLaserLeuValIleGlnValLeuSerAlaValLysTyr 133
Db 446 GCTTTCACGTGAGAGAAGCTGCAGAGATAATGGGGATATTGGCACTGCCATCCAGTTT 505
Qy 134 LeuHisGluAsnGlyIleValHisArgAspLeuLysProGluAsnLeuTyrLeuThr 153
Db 506 CTGCACGCATTAACATTGCCACCGAGATGTCAAGCTTGAAACCTACTCTACATCT 565
Qy 154 ProGluGluAsnSerLysIleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsn 173
Db 566 AAGGAGAAGACAGCAGTGTGTTAAGCTCAGCGATTGTGGCTTGTCTAAGGAGACCCCAA 625
Qy 174 GlyLeuMetSerThrAlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGln 193
Db 626 AATGCCCTGCAGACACCCCTGCTATCTCTCTATATGTGGCCCTGAGGTCTGGGTCCA 685
Qy 194 LysProTyrSerLysAlaValAspCysTyrPsrIleGlyValIleThrTyrIleLeuLeu 213
Db 686 GAGAAGTATGACAAAGTCATGTGACATGTGTCCCTGGGTGCATCATGTACATCCTCTT 745
Qy 214 CysGlyTyrProProPheTyrGluGluThrGluSerLysLeuPhe-----Glu 229
Db 746 TGTGGTCTCCACCTCTACTCCACAGCGGCGCAGGCATCTCCCGGGGATGAAGAGG 805
Qy 230 LysIleLysGluGlyTyrTyrGluPheGluSerProPheTyrAspIleSerGluSer 249
Db 806 AGGATTCGCTGGCGCAGTACGCTCTCCCAATCTCTAGTGTGTCAGAGTCTCTGAGGAT 865
Qy 250 AlaLysAspPheIleCysHisLeuLeuGluLysAspProAsnGluArgTyrThrCysGlu 269
Db 866 GCCAAGCAGCTGATCCGCTCTCTCTTTGTAAGACAGACACCCACAGAGAGGCTGACAT 925
Qy 270 LysAlaLeuSerHisProTyrIleAspGlyAsn-----ThrAlaLeuHis 284
Db 926 CAGTTATGAACCAACCCCTGGATCAACCAATCGATGGTAGTGCCACAGACCCCATCCAC 985
Qy 285 ArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAlaLysSerLysTyr 304
Db 986 -----ACGGCCCGAGTGTCTGCAGAGGAC-----AAGACCACTGG 1021

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-142-551A-1

Alignment Scores:
Pred. No.: 3,46e-43 Length: 1333
Score: 542.50 Matches: 135
Percent Similarity: 52.02% Conservative: 71
Best Local Similarity: 34.09% Mismatches: 139
Query Match: 21.59% Indels: 51
DB: 4 Gaps: 10

US-09-960-643-2 (1-476) x US-09-159-385-3 (1-2132)
Qy 27 GluValLeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGly 46
Db 142 GAGGAGCTGGCGCAGCGCCAGTTTTCGATCGTGGAGAGTCCCGCAGAGAGGCGCGGC 201
Qy 47 LysLeuPheAlaLeuLysCysIleLysLys-----SerProAlaPheArgAspSer 63
Db 202 AAGGAGTACGACGCCCAAGTTTCATCAAGAAGCGCCCGCTGTCTATCCAGCGCGGTGGGTG 261
Qy 64 Ser-----LeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
Db 262 AGCCGGGAGGAGATCGACGGGAGGTGAACATCTCTGGCGGAGATCCCGCACCCCAACATC 321
Qy 81 ValThrLeuGluAspIleTyrGluSerThrHisTyrTyrLeuValMetGlnLeuVal 100
Db 322 ATCACCTTCGACGACATCTTCGAGAACAGACGCGGCTGGTCCATCCTGGAGCTGTC 381
Qy 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
Db 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
```

Db	382	TC	TGCGGGGAGCTCTTTGACTTCCTGCGGGAGAAAGAGTCGCTGACGGAGCAGGACC	441
Qy	121	Ser	LeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal	140
Db	442	ACC	CAGTTCCTCAAGCAGATCCTGGAGCGGTCTACCTACCTGCATCTTAAGCGCATCGCA	501
Qy	141	His	ArgAspLeuLysProGluAsnLeuTyrLeuThrProGlu---GluAsnSerLys	159
Db	502	CAC	TTCACCTGAAGCCGGAACACATCATGCTGCTGGACAAGAAGCTGCCCAACCCACGA	561
Qy	160	Ile	MetIleThrAspPheGlyLeuSer---LysMetGluGlnAsnGlyIleMetSerThr	178
Db	562	ATCA	AGCTCATCATCGCATCGGCACACAGATCGAGCGGGGAACGAGTTCAAGAAC	621
Qy	179	Ala	CysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys	198
Db	622	ATCT	TGGCACCCCGAGTGTGTGCCCCAGAGATGTGAATGACCGCTGGGCGTG	681
Qy	199	Ala	ValAspCysTyrSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro	218
Db	682	GAG	CGGACATGTGGAGCATCGGTGCATCACCATTATCTCTCTGAGCGGTGCATCCCG	741
Qy	219	Phe	TyrGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe	238
Db	742	TTCT	TGGCGAGACCAAGCAGGAGCGCTCAACCAACATCTCAGCGCTGAACCTAGCATTC	801
Qy	239	Glu	SerProPheTyrAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu	258
Db	802	GAG	CAGGAGTACTTTCACACACCGAGCGAGCTGGCCAGGACTTCATTTCGCCGCGCTGC	861
Qy	259	Glu	LysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTyrPileAsp	278
Db	862	GTCA	AGATGCCAAGCGGAGATGACCATTCGCCAGAGCCTGGAAACATTCCTGGATTAA-	920
Qy	279	Gly	AsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn	298
Db	921	GG	CGATCCCGCG-----	932
Qy	299	Phe	AlaLysSerLysTyrPArgGlnAlaPheAsnAlaAlaValValHisHisMetArg	318
Db	933	---	CGCGAACGTGCGTG-----	947
Qy	319	Lys	LeuHisMetAsnLeuHisSerProGlyValArgProGluValGluValAsnArgProPro	338
Db	948	-----	-----TCAGCAGCGCGCGCGACCCGCGCGCGCGCT	983
Qy	339	Glu	ThrGlnAlaSerGluThrSerArgProSerProGluIleThrIleThrGluAla	358
Db	984	GAAG	CACGCGTCTGAAGG-AGTACACCATCAAGT-----CGCACT	1024
Qy	359	Pro	ValLeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuPro---CysGlnHis	377
Db	1025	CCAG	CTGCGGCCCAACACAGCTACGCGACTTCGAGCGCTCTCTCCAAAGTGTGTGGAGG	1084
Qy	378	Gly	ArgArgProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGly-SerLe	397
Db	1085	AG	CGCGGCGCG-----CCGAGGAGGCGCTCGCGAGCTCGAGCGCAGCGCGCGCTCT	1138
Qy	397	u	-----HisIle-----SerSerSerLeuValProMet	406
Db	1139	GCC	ACGAGACGTGGAGGCGCTGGCGGCATCTACGAGGAGAGGAGGCGCTGGTACCG--	1196
Qy	406	thi	GlnGlySerLeuAlaAlaGlyProCysGlyCysSerSerCysLeuAsnIleGln	426
Db	1197	----	CGAGGAGCGACAGCCTGGGCCAGG-ACCTGC-----GG	1230
Qy	426	yse	LysGlyLysSerSerTyrCysSerGluProThrLeuLeuLysLysAlaAsnLysLys	446
Db	1231	AG	GCTACGCGCAGGAGC-----TGCTCAAGACCGAGG---CGCTCAACGCGCAGCGCAG	1281
Qy	446	sgl	AsnPheLysSerGluValMetValProValLysAlaSerGlySerSer	463
Db	1282	GAG	AGGCCCAAGGGCGCGTGTGGGAGCAGCGGCTTCAAGCGCGCTTCA	1333

```

RESULT 12
US-09-186-277-3
: Sequence 3, Application US/09186277
: Patent No. 6171841
: GENERAL INFORMATION:
: APPLICANT: AKIRA, SHIZUO
: APPLICANT: KAWAI, TARO
: TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
: FILE REFERENCE: 081356/0128
: CURRENT APPLICATION NUMBER: US/09/186,277
: CURRENT FILING DATE: 1998-11-05
: EARLIER APPLICATION NUMBER: JP97/261589
: EARLIER FILING DATE: 1997-09-26
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3
: LENGTH: 2132
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (94)..(1455)
US-09-186-277-3

```

US-09-960-643-2 (1-476) x US-09-186-277-3 (1-2132)

Qy	27	GluValLeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGly	46
Db	142	GAGAGCTGGGACCGCCAGTTTGGCATCTGCGGAAGTGC CGGCAAGGACGGCCAGCGGC	201
Qy	47	LysLeuPheAlaLeuLysCysIleLysLys-----SerProAlaPheArgAspSer	63
Db	202	AAGGAGTACGCAGCCAGTTTCATCAAGAAGCGCGCTGTCATCCAGCGCGCTGGGGTG	261
Qy	64	Ser-----LeuGluAsnGluIleAlaValLeuLysLysLysHisGluAsnIle	80
Db	262	AGCGCGGAGGAGATCGACGGGAGGTGAACATCCTCGGAGATCGGCACCCCAACATC	321
Qy	81	ValThrLeuGluAspIleTyrGluSerThrHisTyrThrLeuValMetGlnLeuVal	100
Db	322	ATCACCTGTCAGACATCTTCGAGAACACACGCGAGCTGGTCTCATCTCGGAGCTGGTC	381
Qy	101	SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla	120
Db	382	TCTGCGGGAGCTCTTTGACTTCTCGCGGAGAAAGATCGCTGACGGAGGACGAGGCC	441
Qy	121	SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal	140
Db	442	ACCCAGTTCCTCAAGCAGATCTCGAGCGGCTTCACTACCTGCATCTAAGCGCATCGCA	501
Qy	141	HisArgAspLeuLysProGluAsnLeuTyrLeuThrProGlu---GluAsnSerLys	159
Db	502	CACTTTCACCTGAGCGGAAACATCATGCTGCTGGCAAGAAAGCTGCCCAACCCACGA	561
Qy	160	IleMetIleThrAspPheGlyLeuSer---LysMetGluGlnAsnGlyIleMetSerThr	178
Db	562	ATCAAGTTCATCGACTTCGGCATCTCGCCACAAGATCCAGGCGGGGAACGAGTTCAAGAAC	621
Qy	179	AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys	198
Db	622	ATCTTCGGCACCCCGGAGTTTGTGGCCCCAGAGATTGTGAACATGAGCGCGCTGGGCCTG	681
Qy	199	AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro	218

```
Db 682 GAGCGGACATGTGGAGCATCGTGTATCATACCTATATCTCTGAGCGGTGCATCCCG 741
Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysGluGlyTyrTyrGluPhe 238
Db 742 TTCCTGGCGGACCAAGCAGGAGCGCTCACACATCTCAGCGTGAACTACGACATC 801
Qy 239 GluSerProPheTrpAspPheSerGluSerAlaLysAspPheLysHisLeuLeu 258
Db 802 GACGAGGAGTACTTACGACACACGAGCTGGCCAGGACTTCATTCGCGGCTGCTC 861
Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpLeasp 278
Db 862 GTCAAGATCCCAAGCGGAGATGACCATTTGCCAGAGCTGGAACATTCCTGGATTAA- 920
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
Db 921 GCGGATCCGCGC----- 932
Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValHisMetArg 318
Db 933 ---GCGGACGTCGTGG----- 947
Qy 319 LysLeuHisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProPro 338
Db 948 -----TGAGGACAGCGCGCCCAAGCCGCGAGCGCGCGCGCT 983
Qy 339 GluThrGlnAlaSerGluThrSerArgProSerSerProGluIleThrGluAla 358
Db 984 GAAGACACCGCTCTGAAGG-AGTACACCATCAAGT-----CGCACT 1024
Qy 359 ProValLeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuPro---CysGlnHis 377
Db 1025 CCAGCTTGGCGGCCCAACAGACTACGCCGACTTCGAGCGCTTCTCAAGGTGCTGGAGG 1084
Qy 378 GlyArgArgProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGly-SerLe 397
Db 1085 AGCGCGCGCGC-----CCGAGGAGGCGCTGCGGAGCTGAGCGAGCGAGCGCGGCTCT 1138
Qy 397 u-----Hisile-----SerSerSerLeuValProMe 406
Db 1139 GCACAGGAGCGTGAGCGCGCTGGCGCCCTCTACGAGGAGGAGGAGGCGCTGTACCG-- 1196
Qy 406 tHisGlnGlySerLeuAlaAlaGlyProCysGlyCysCysSerSerCysLeuAsnIleGl 426
Db 1197 ----CGAGGAGAGCGAGCGCTGGCGCAGG-ACCTGC-----GG 1230
Qy 426 ySerLysGlyLysSerSerTyrCysSerGluProThrLeuLeuLysLysAlaAsnLysLy 446
Db 1231 AGGCTACGCGGAGGAGC-----TGCTCAAGACCGAGG---CGCTCAAGCGGAGCGCGCAG 1281
Qy 446 sGlnAsnPheLysSerGluValMetValProValLysAlaSerGlySerSer 463
Db 1282 GAGGAGCCCAAGGCGCGCTGCTGGGAGCAGCGCGCTCAAGCGCGCTCA 1333

RESULT 13
US-08-464-164-1
; Sequence 1, Application US/08464164
; Patent No. 5614195
; GENERAL INFORMATION:
; APPLICANT: Tonley, Fiona M.
; APPLICANT: Dunn, Paul P. J.
; APPLICANT: Bumstead, Janene M.
; APPLICANT: Vermeulen, Arno N.
; TITLE OF INVENTION: Coccidiosis poultry vaccine
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5614195el patent Department
; STREET: 1300 Piccard Drive, Sulte 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/464,164
; APPLICATION NUMBER: US/08/464,164
; FILING DATE: June 2, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: C-terminal
; ORIGINAL SOURCE:
; ORGANISM: Eimeria maxima
; STRAIN: Houghton
; DEVELOPMENTAL STAGE: sporozoite
; IMMEDIATE SOURCE:
; LIBRARY: sporozoite cDNA cloned in Lambda ZAPII
; CLONE: Em70-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1368
; US-08-464-164-1

Alignment Scores:
Pred. No.: 1,14e-42 Length: 1400
Score: 537.50 Matches: 122
Percent Similarity: 57.70% Conservative: 54
Best Local Similarity: 40.00% Mismatches: 116
Query Match: 21.39% Indels: 13
DB: 6

US-09-960-643-2 (1-476) x US-08-464-164-1 (1-1400)
Qy 25 PheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnArgLeu 44
Db 4 TTCGTGGAAGTTTGGGT-----GAGGTCTCTTATGCAAGGACAAGATA 48
Qy 45 ThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPheArgAspSerSer 64
Db 49 ACAGCACAGGAATATGCAATAAGTAATATCTAAACGCTCAAGTAAACACAGAAGACAGAT 108
Qy 65 LeuGlu-----AsnGluIleAlaValLeuLysLysIleLysHisGluAsnIleVal 81
Db 109 AAAGAATTATTTAAAGAGAGTTGAATTATTAAGAATAATTAGATCATCTTAATATCATG 168
Qy 82 ThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuValSer 101
Db 169 AAATTATATGAATCTTTGAGGATAAAGGATACTTTTATCTTTGTTACAGAAAGTATATACA 228
Qy 102 GlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAlaSer 121
Db 229 GGAGGAGAATTATTTGATGAAATTAATTAACGAAAGATTTCAGGAGCGCGGATGCAGCT 288
Qy 122 LeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleValHis 141
Db 289 CGTATAGTACGTCAGGTTCTATCGGTTATTAATTATATGTCATCGTAATAAATATGTTTCA 348
Qy 142 ArgAspLeuLysProGluAsnLeuLeuTyrThrProGluGluAsnSerLysIleMet 161
Db 349 AGAGATTTAAAGCCAGAGAATTATTTATTATAGAGAATAAAAAAAGATGCAATATACGA 408
```









GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:04:21 ; Search time 14 Seconds  
(without alignments)  
1410.195 Million cell updates/sec

Title: US-09-960-643-2

Perfect score: 2513

Sequence: 1 MGRKEEDDCSSWKKQTNR.....VKASGSSHCRAQTGVCCLIM 476

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1155.5	46.0	374	1 KCC1_RAT	Q63450 rattus norv
2	1152.5	45.9	370	1 KCC1_HUMAN	Q14012 homo sapien
3	697.5	27.8	469	1 KCCA_MOUSE	P08414 mus musculu
4	694.5	27.6	295	1 KMLC_DICDI	P25323 dictyosteli
5	692.5	27.6	474	1 KCC4_RAT	P13234 rattus norv
6	684	27.2	473	1 KCC4_HUMAN	Q16566 homo sapien
7	651	25.9	414	1 KCC1_EMEFI	Q00771 emericella
8	643.5	25.6	424	1 KPSH_HUMAN	P11801 homo sapien
9	630	25.1	335	1 KCC1_SCHPO	Q9P712 schizosacch
10	612	24.4	740	1 DCK1_HUMAN	O15075 homo sapien
11	606	24.1	756	1 DCK1_MOUSE	Q9J1M8 mus musculu
12	603.5	24.0	664	1 KCCB_HUMAN	Q13554 homo sapien
13	602.5	24.0	433	1 DCK1_RAT	O08875 rattus norv
14	602	24.0	478	1 KCCA_HUMAN	Q9UCM7 homo sapien
15	599	23.8	478	1 KCCA_RAT	P11275 rattus norv
16	597	23.8	546	1 CHK2_MOUSE	Q9Z265 mus musculu
17	594.5	23.7	499	1 KCCD_HUMAN	Q13557 homo sapien
18	591.5	23.5	533	1 KCCD_RAT	P15791 rattus norv
19	588	23.4	447	1 KCC2_YEAST	P22517 saccharomyc
20	584	23.2	478	1 KCCA_MOUSE	P11798 mus musculu
21	579.5	23.1	542	1 KCCB_MOUSE	P28652 mus musculu
22	578.5	23.0	542	1 KCCB_RAT	P08413 rattus norv
23	578.5	23.0	543	1 CHK2_HUMAN	O96017 homo sapien
24	578	23.0	472	1 KCCG_HUMAN	Q13555 homo sapien
25	577.5	23.0	446	1 KCC1_YEAST	P27466 saccharomyc
26	574	22.8	529	1 KCCG_MOUSE	Q923T9 mus musculu
27	566.5	22.5	527	1 KCCG_RAT	P11730 rattus norv
28	565.5	22.5	382	1 KCC1_METAN	O14408 metarhizium
29	558	22.2	821	1 SPK1_YEAST	P22216 saccharomyc
30	556.5	22.1	610	1 CDPI_ARATH	Q06850 arabidopsis
31	543.5	21.6	508	1 CDPK_SOYBN	P28583 glycine max
32	543.5	21.6	1431	1 DAPK_HUMAN	P53355 homo sapien
33	536	21.3	512	1 RCK1_YEAST	P38622 saccharomyc

RESULT 1

KCC1\_RAT

ID KCC1\_RAT STANDARD; PRT; 374 AA.  
AC Q63450; Q63084;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Calcium/calmodulin-dependent protein kinase type I (EC 2.7.1.123)  
DE (CAM kinase I).  
GN CAMK1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP MEDLINE=94075341; PubMed=8253780;  
RA Piccotto M.R., Czernik A.J., Nairn A.C.;  
RT "Calcium/calmodulin-dependent protein kinase I. cDNA cloning and  
RT identification of autophosphorylation site";  
RL J. Biol. Chem. 268:26512-26521(1993).  
RN [2]  
RP SEQUENCE FROM N.A. (42 KDA ISOFORM).  
RX STRAIN=Sprague-Dawley; TISSUE=Lung;  
RC MEDLINE=95035115; PubMed=7948038;  
RA Cho F.S., Phillips K.S., Boquetti B., Weaver T.E.;  
RT "Characterization of a rat cDNA clone encoding calcium/calmodulin-  
RT dependent protein kinase I.";  
RL Biochim. Biophys. Acta 1224:156-160(1994).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF SHORT FORM.  
RC TISSUE=Brain;  
RX MEDLINE=96182648; PubMed=8601311;  
RA Goldberg J., Nairn A.C., Kuriyan J.;  
RT "Structural basis for the autoinhibition of  
RT calcium/calmodulin-dependent protein kinase I.";  
RL Cell 84:875-887(1996).  
CC -1- FUNCTION: PHOSPHORYLATES SYNAPSIN I.  
CC -1- CATALYTIC ACTIVITY: ATP + protein = ADP + O-phosphoprotein.  
CC -1- ENZYME REGULATION: ACTIVATED BY CA++/CALMODULIN. MUST BE  
CC -1- PHOSPHORYLATED TO BE MAXIMALLY ACTIVE.  
CC -1- SUBUNIT: MONOMER.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC CAMK SUBFAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; L24907; AAA19670.1; -;  
DR EMBL; L26288; AAA66944.1; -;  
PDB; 1A06; 08-APR-98.

34 536 21.3 533 1 CDP2\_ORYSA  
35 531 21.1 735 1 KGA1\_HUMAN  
36 531 21.1 1176 1 KMLS\_POVIN  
37 530.5 21.1 735 1 KGA1\_RAT  
38 530 21.1 1147 1 KMLS\_RABIT  
39 529 21.1 513 1 CDP2\_MAI2E  
40 527 21.0 724 1 KGA1\_MOUSE  
41 526.5 21.0 542 1 CDP3\_ORYSA  
42 525 20.9 1906 1 KMLS\_CHICK  
43 523 20.8 513 1 DUN1\_YEAST  
44 521 20.7 1914 1 KMLS\_HUMAN  
45 520 20.7 534 1 CDP1\_ORYSA

ALIGNMENTS

```
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW Phosphorylation; ATP-binding; Alternative splicing; 3D-structure.
FT DOMAIN 20 276
FT DOMAIN 287 321
FT NP_BIND 26 34
FT BINDING 49 49
FT ACT_SITE 141 141
FT MOD_RES 177 177
FT VARSPLIC 323 374
FT PAPPSSRAMD -> HOPGGTGTDG (IN 37 KDA ISOFORM).
FT F -> G (IN REF. 1).
FT CONFLICT 112 112
FT CONFLICT 118 118
FT CONFLICT 309 309
FT SEQUENCE 374 AA; 41638 MW; 37889B3DEF033AB2 CRC64;

Query Match 46.0%; Score 1155.5; DB 1; Length 374;
Best Local Similarity 69.6%; Pred. No. 1.3e-65;
Matches 222; Conservative 43; Mismatches 51; Indels 3; Gaps 3;

QY 12 WKQTTNIRKTFIPMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPAFRDSLSLENEIA 70
DB 10 W-KQAEIDIRYDFRDVLGTGAFSEVFLAEDKRTQKLVAKICAKALEGKSGMENEIA 68
QY 71 VLKIKHENIVTLEDIYESTHYLYVMQLVSGGELFDRILRGVYTERKDSLVIQQLVLSA 130
DB 69 VLHKIKHPNIVALDDIYESSGHLHYLMQLVSGGELFDRIVEKGFYTERDASRLIFQVLDA 128
QY 131 VKYLHENGIVHRDLKPNLLYLTPEENSKIMITDFGLSKMEQNG-INSTACGTPGYVAPE 189
DB 129 VKYLDHGLGVHRDLKPNLLYSLDESKIMISDFGLSKMEDPGSVLSTACGTPGYVAPE 188
QY 190 VLAQPKYSKAVDCWSIGVITYILLCGYPFPYEEETESKLFKEKIGYEFESFPWDDISDS 249
DB 189 VLAQPKYSKAVDCWSIGVITYILLCGYPFPYDENDAKLFEQILKAIEYFDSFYWDDISDS 248
QY 250 AKDFICHLLEKDPNERYTCEKALSHPWIDGNLTALHRDIYPSVSLQIQKNFAKSKWRQAFN 309
DB 249 AKDFIRHLMKDEPKRFTCEQALQHPWAGTDALDKNIHQSVSEQIKKFNFAKSKWKQAFN 308
QY 310 AAAYVHHMRKLHMLHSPG 328
DB 309 ATAVVRHMRKLQLTSGEQ 327

RESULT 2
KCCL_HUMAN STANDARD; PRT; 370 AA.
AC Q14012;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcium/calmodulin-dependent protein kinase type I (EC 2.7.1.123)
DE (CAM kinase I).
GN CAMK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95369239; PubMed=7641687;
RA Haribabu B., Hook S.S., Seibert M.A., Goldstein E.G., Tomhave E.D.,
RA Edelman A.M., Snyderman R., Means A.R.;
RT "Human calcium-calmodulin dependent protein kinase I: cDNA cloning,
```

```
RT domain structure and activation by phosphorylation at threonine-177
RT by calcium-calmodulin dependent protein kinase I kinase.";
RL EMBO J. 14:3679-3686(1995).
CC 1- FUNCTION: PHOSPHORYLATES SYNAPSIN I.
CC 1- CATALYTIC ACTIVITY: ATP + protein -> ADP + O-phosphoprotein.
CC 1- ENZYME REGULATION: ACTIVATED BY CA++/CALMODULIN. MUST BE
CC PHOSPHORYLATED TO BE MAXIMALLY ACTIVE.
CC 1- SUBUNIT: MONOMER.
CC 1- TISSUE SPECIFICITY: UBQUITOUS.
CC 1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CAMK SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L41816; AAA99458.1; ..
CC HSSP; Q63450; IA06..
CC Genew; HGNC:1459; CAMK1.
CC MIN; 604998; ..
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR002290; Ser_thr_pkinase.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Euk_pkinase; 1.
CC SMART: SM00220; S_TKC; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW Phosphorylation; ATP-binding; Alternative splicing.
FT DOMAIN 20 276
FT DOMAIN 287 321
FT NP_BIND 26 34
FT BINDING 49 49
FT ACT_SITE 141 141
FT MOD_RES 177 177
FT MUTAGEN 49 49 K->A: LOSS OF ACTIVITY.
FT SEQUENCE 370 AA; 41337 MW; 57FA20BCE00FA76C CRC64;

Query Match 45.9%; Score 1152.5; DB 1; Length 370;
Best Local Similarity 55.8%; Pred. No. 2e-65;
Matches 239; Conservative 46; Mismatches 62; Indels 81; Gaps 9;

QY 12 WKQTTNIRKTFIPMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPAFRDSLSLENEIA 70
DB 10 W-KQAEIDIRYDFRDVLGTGAFSEVFLAEDKRTQKLVAKICAKALEGKSGMENEIA 68
QY 71 VLKIKHENIVTLEDIYESTHYLYVMQLVSGGELFDRILRGVYTERKDSLVIQQLVLSA 130
DB 69 VLHKIKHPNIVALDDIYESSGHLHYLMQLVSGGELFDRIVEKGFYTERDASRLIFQVLDA 128
QY 131 VKYLHENGIVHRDLKPNLLYLTPEENSKIMITDFGLSKMEQNG-IMSTACGTPGYVAPE 189
DB 129 VKYLHGLGVHRDLKPNLLYSLDESKIMISDFGLSKMEDPGSVLSTACGTPGYVAPE 188
QY 190 VLAQPKYSKAVDCWSIGVITYILLCGYPFPYEEETESKLFKEKIGYEFESFPWDDISDS 249
DB 189 VLAQPKYSKAVDCWSIGVITYILLCGYPFPYDENDAKLFEQILKAIEYFDSFYWDDISDS 248
QY 250 AKDFICHLLEKDPNERYTCEKALSHPWIDGNLTALHRDIYPSVSLQIQKNFAKSKWRQAFN 309
DB 249 AKDFIRHLMKDEPKRFTCEQALQHPWAGTDALDKNIHQSVSEQIKKFNFAKSKWKQAFN 308
QY 310 AAAYVHHMRKLHMLHSPG 328
DB 309 ATAVVRHMRKLQLTSGEQ 327
QY 370 LTQLPCQGRPRTPAGGRSLNCLVNGSLHSSSLVPMHQSLAAGPCGCCSCLNIGSKG 429
```

```

Db 325 -----QEQGQTAFSHGELLTPVAGGP-----AAG-C-CRDC----- 354
QY 430 KSSVCSEP 437
Db 355 -----CVEP 358

RESULT 3
KCC4_MOUSE STANDARD; PRT; 469 AA.
AC P08414; Q61381;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calcium/calmodulin-dependent protein kinase type IV catalytic chain
DE (SC 2.7.1.123) (CAM kinase-GR) (CaMK IV).
GN CAMK4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=91372388; PubMed=1893997;
RA Jones D.A., Glod J., Wilson-Shaw D., Hahn W.E., Sikela J.M.;
RT "cDNA sequence and differential expression of the mouse
RT Ca2+/calmodulin-dependent protein kinase IV gene.";
RL FEBS Lett. 289:105-109(1991).
RN [2]
RP SEQUENCE OF 240-469 FROM N.A.
RX MEDLINE=89122027; PubMed=2536634;
RA Sikela J.M., Law M.L., Kao F.-T., Hartz J.A., Wei Q., Hahn W.E.;
RT "Chromosomal localization of the human gene for brain
RT Ca2+/calmodulin-dependent protein kinase type IV.";
RL Genomics 4:21-27(1989).
RN [3]
RP SEQUENCE OF 315-469 FROM N.A.
RX MEDLINE=87204263; PubMed=3033675;
RA Sikela J.M., Hahn W.E.;
RT "Screening an expression library with a ligand probe: isolation and
RT sequence of a cDNA corresponding to a brain calmodulin-binding
RT protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:3038-3042(1987).
CC -1- CATALYTIC ACTIVITY: ATP + protein = ADP + O-phosphoprotein.
CC -1- SUBCELLULAR LOCATION: SUBSTANTIAL LOCALIZATION IN CERTAIN NEURONAL
CC NUCLEI.
CC -1- TISSUE SPECIFICITY: BRAIN AND TESTIS.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CAMK SUBFAMILY.

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M16206; AAA39933.1; -
DR EMBL; M64266; AAA37364.1; -
DR EMBL; J03057; AAA37366.1; -
DR EMBL; X58995; CAA41741.1; -
DR PIR; A29878; A29878.
DR PIR; S17656; S17656.
DR HSSP; Q63450; 1A06.
DR MGD; MGI:88258; Camk4.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.

```

```

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Calmodulin-binding.
FT DOMAIN 42 296 PROTEIN_KINASE.
FT NP_BIND 48 56 ATP (BY SIMILARITY).
FT BINDING 71 71 ATP (BY SIMILARITY).
FT ACT_SITE 160 160 BY SIMILARITY.
FT DOMAIN 318 337 CALMODULIN-BINDING (POTENTIAL).
FT CONFLICT 278 280 VLD -> CFGI (IN REF. 2).
FT CONFLICT 302 302 N -> T (IN REF. 2).
SQ SEQUENCE 469 AA; 52627 MW; CE1F98670822F975 CRC64;

Query Match 27.8%; Score 697.5; DB 1; Length 469;
Best Local Similarity 42.7%; Pred. No. 8.6e-37;
Matches 152; Conservative 60; Mismatches 121; Indels 23; Gaps 7;

QY 9 CSSMKKQTTNTRKTF-----IFMEV-----LGSGAFSEVFLVKORLTGKLFALKCI 54
DB 14 CSSTASTENLVDPYWDGSDNRDPLGDFEVESELGRGATSVIRCKQKQTKPYALKVL 73
QY 55 KKSFAFRDSSLENEITAVLKIKHENIVTLEDIYESTTHYLYVMOLVSGGELFDRILRGV 114
DB 74 KKT--VDKKIVRTEIGVLLRLSHPNIIKLEIFETPTTEISLVLELVGTGGELFDRIVEKG 131
QY 115 YTEKASLVIOOVLNAVYKYLHENGIVHRDLKAPENLLYLTPEENSKIMITDFGLSK-MEQN 173
DB 132 YSERDARDAVQKILEAVAYLHENGIVHRDLKAPENLLYATPAPADAPLKIADFGLSKIYEHQ 191
QY 174 GIMSTACGTPGVVAPVLAQPKYKAVDQWSIGVITYILLGYPFPVEET--ESKLFPEIK 232
DB 192 VLMKTVCTPGYCAPEILRGCAYPEVDMVSGVIGIYILLCGFEFFDDEGDDQFMRRL 251
QY 233 EGYEFESFPWDDISESAKDFICHLLEKDPNERYTCERKALSHPWIDGNTA--LHRDIYPS 290
DB 252 NCEYFIFSPWDEVSNAKDLVKLLVLDPKKRLTTTQALQHPWVTGKAANFVHMD---T 308
QY 291 VSLQTKNFAKSKWFOAFNAAVHHMKLHMNLHSGVREVENRRPPEQASSETS 346
DB 309 AQKQLQEFNARRKKAQKAVVASSRSGSSSHTSIQENHKASSDPSPQDAKDS 364

RESULT 4
KMLC_DICDI STANDARD; PRT; 295 AA.
AC P25323;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin light chain kinase (EC 2.7.1.17) (MLCK).
GN MLKA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=91340753; PubMed=1651931;
RA Tan J.L., Spudich J.A.;
RT "Characterization and bacterial expression of the Dictyostelium
RT myosin light chain kinase cDNA. Identification of an autoinhibitory
RL domain.";
RL J. Biol. Chem. 266:16044-16049(1991).
RN [2]
RP REVISIONS.
RA Spudich J.A.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP PARTIAL SEQUENCE.
RC STRAIN=AX3;
RX MEDLINE=90337997; PubMed=2380188;
RA Tan J.L., Spudich J.A.;

```



```
DR EMBL; M63334; AAA40865.1; -.
DR EMBL; M74488; AAA40845.1; ALT_SEQ.
DR EMBL; M64757; AAA40856.1; -.
DR EMBL; M64757; AAA40857.1; -.
DR EMBL; J04600; AAA41867.1; -.
DR EMBL; J04446; AAA40990.1; -.
DR PIR; A41103; TVRTC4.
DR HSP; Q63450; 1A06.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Calmodulin-binding; Testis; Alternative splicing.
FT DOMAIN 42 296 PROTEIN KINASE.
FT NP_BIND 48 56 ATP (BY SIMILARITY).
FT BINDING 71 71 ATP (BY SIMILARITY).
FT ACT_SITE 160 160 BY SIMILARITY.
FT DOMAIN 318 337 CALMODULIN-BINDING (POTENTIAL).
FT DOMAIN 393 399 POLY-GLU.
FT DOMAIN 403 413 POLY-GLU.
FT VARSPIC 1 305 MISSING (IN ISOFORM 2).
FT CONFLICT 372 372 I -> M (IN REF. 2 AND 4).
SQ SEQUENCE 474 AA; 53133 MW; 56f71AC5644DED23 CRC64;

Query Match 27.6%; Score 692.5; DB 1; Length 474;
Best Local Similarity 40.8%; Pred. No. 1.8e-36;
Matches 158; Conservative 62; Mismatches 120; Indels 47; Gaps 10;

QY 9 CSWKKQTNIKTF-----IPMEV---LGSGAFSEVFLVKQLTGKLFALKCI 54
Db 14 CSVTSSTENLPDYWDIDGSKRDLSDFFVESELSGRGATSIYRCQKQKPYALKVL 73
QY 55 KKSAPFRDSSLENAVLKIKHENIVTLEDIYESTHYLVLMQVSGSELPDRLTGRV 114
Db 74 KKT--VDKKIVTEIGVLLRSHPNLIIKLEFETPTISLVLEVTGELFDRIVEAGY 131
QY 115 YTEKDALVIOQVLSAVKYLHENGIVHRDLKPENLLYLTPENSKIMITDFGLSK-MEQN 173
Db 132 YSERDAADAVKQILEAVAYLHENGIVHRDLKPENLLYATPADAPLKIADFGLSKIVEHQ 191
QY 174 GIMSTACGTPGVYAVEVLAQKPYSKAVDCWSIGVTYILLCGYPFFYEET-ESKLFERIK 232
Db 192 VLMKTVCGPGPGYCAPEILRGCAYGPEVDMWSVGIITYILLCGFEPFYDERGDFMFRIL 251
QY 233 EGYEYFESPFWDDISESARDFTCHLEKDPNERYTCEKALSHPWIDGNTA--LHRDIYPS 290
Db 252 NCEYFYFISPWDEVSINAKDLVKLLVLDPKKRLTTFQALQHPWVTGKAANFVHMD---T 308
QY 291 VSLQTKNPAKSKWRQAFNA-----AAVVHMRKRLMNLHSPGVRP-----EV 333
Db 309 AQKLLQEFNARKRLKAAYKAVVASSRLGSASSSHTNIQESNKASSEAPQADGDKTDPL 368
QY 334 ENR-----PPETOASETSRSPEI 353
Db 369 ENKIQAGDHEAAKAAADETMKLGQSEV 395

RESULT 6
KCC4_HUMAN
ID KCC4_HUMAN STANDARD; PRT; 473 AA.
AC Q16566;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calcium/calmodulin-dependent protein kinase type IV catalytic chain
DE (EC 2.7.1.123) (CaM kinase-GR) (CaMK IV).
GN CAMK4.
OS Homo sapiens (Human).
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. [1] SEQUENCE FROM N.A. MEDLINE=94375404; PubMed=8089075; Kitani T., Okuno S., Fujisawa H.; "CDNA cloning and expression of human calmodulin-dependent protein kinase IV."; J. Biochem. 115:637-640(1994). [2] SEQUENCE FROM N.A. TISSUE=Cerebellum, and Thymus; MEDLINE=94252566; PubMed=8194751; Bland M.M., Monroe R.S., Ohmstede C.A.; "The cDNA sequence and characterization of the Ca2+/calmodulin-dependent protein kinase-Gr from human brain and thymus."; Gene 142:191-197(1994). [3] SEQUENCE FROM N.A. TISSUE=Blood; MEDLINE=94149862; PubMed=8107230; Mosialos G., Hanissian S.H., Jawaahar S., Vara L., Kieff E., Chatila T.A.; "A Ca2+/calmodulin-dependent protein kinase, Cam kinase-Gr, expressed after transformation of primary human B lymphocytes by Epstein-Barr virus (EBV) is induced by the EBV oncogene LMP1."; J. Virol. 68:1697-1705(1994). [4] SEQUENCE FROM N.A. TISSUE=Brain; Strausberg R.; Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases. -1- FUNCTION: CALSPERMIN IS A HEAT-STABLE, ACIDIC, CALMODULIN-BINDING PROTEIN. -1- CATALYTIC ACTIVITY: ATP + protein = ADP + O-phosphoprotein. -1- SUBCELLULAR LOCATION: SUBSTANTIAL LOCALIZATION IN CERTAIN NEURONAL NUCLEI (BY SIMILARITY). -1- ALTERNATIVE PRODUCTS: CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE TYPE IV CATALYTIC CHAIN AND CALSPERMIN ARE OBTAINED BY ALTERNATIVE SPLICING. -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. ----- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). ----- EMBL; D30742; BAA06403.1; -. EMBL; L17000; AAA35639.1; -. EMBL; L24959; AAA18251.1; -. EMBL; BC016695; AAH16695.1; -. HSSP; Q63450; 1A06. Genew; HGNC:1464; CAMK4. MIM; 114080; -. InterPro; IPR000719; Euk\_pkinase. InterPro; IPR002290; Ser\_thr\_pkinase. Pfam; PF00069; pkinase; 1. ProDom; PD000001; Euk\_pkinase; 1. SMART; SM00220; S\_TKC; 1. PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1. PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1. PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1. Transferrase; Serine/threonine-protein kinase; ATP-binding; Calmodulin-binding. DOMAIN 46 300 PROTEIN KINASE. FT NP\_BIND 52 60 ATP (BY SIMILARITY). FT BINDING 75 75 ATP (BY SIMILARITY).

```
FT ACT_SITE 164 341 BY SIMILARITY.
FT DOMAIN 322 341 CALMODULIN-BINDING (POTENTIAL).
SQ SEQUENCE 473 AA; 51925 MW; 8FEE51E5612326DC CRC64;

Query Match 27.2%; Score 684; DB 1; Length 473;
Best Local Similarity 48.1%; Pred. No. 6.1e-36;
Matches 141; Conservative 53; Mismatches 87; Indels 12; Gaps 6;

Qy 25 FMEV---LGSAFSEVFLVKQRTGKLFALKCIKKSPAFRDSLSLENEIAVLKKIHENIV 81
Db 45 FFEVESELGRGATSVIVRCQKQKQKPYALKVLKKT--VDKKIVRTGIVGLRLSHPNII 102
Qy 82 TLEDIYSTTHYLLVMQLVSGGELFDRILRGVYTEKDSALVIOQVLSAVKYLVHENGIVH 141
Dy 103 KLKEIFETPTISLVLELTGGLFDRIVEKGYYSERDAADAVKQILEAVAYLHENGIVH 162
Qy 142 RDLKPENLLYTPENSKIMTIDGLSK-MEONGIMSTAGTGPYVAPEVLAQPKYSKAV 200
Dy 163 RDLKPENLLYATPADAPLKIADFLGSLVHEQVLMKTVCGTGYCAPEILRGCAYPEV 222
Qy 201 DCWSIGVITYLLCGYPPFYEEET-ESKLFEEKIKEGYEFESPFWDIDISESAKDFICHLLE 259
Dy 223 DMWSGIITYLLCGFPFEDRGDQFMRRLNCEYFFISPWDEVSINAKDLVRKLIIV 282
Qy 260 KDPNRYTCERKALSHPWIDGNTA--LHRDIYPSVLSIQKNFAKSKWRQAFNA 310
Dy 283 LDPKKRLTTFQALQHPWVTGKAANFVHMD---TAQKKLQEFNARRKLKA AVKA 332

RESULT 7
KCC1_EMENI STANDARD; PRT; 414 AA.
AC Q00771;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calcium/calmodulin-dependent protein kinase (EC 2.7.1.123) (CMPK).
GN CMKA.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eutiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9225350; PubMed=1563634;
RA Kornstein L.B., Gaiso M.L., Hammell R.L., Bartelt D.C.;
RT "Cloning and sequence determination of a cDNA encoding Aspergillus
nidelans calmodulin-dependent multifunctional protein kinase.";
RL Gene 113:75-82(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=R153;
RA Subbaramiah K., Greene V., Bartelt D.C.;
RT "Structure of the cmka gene encoding a CamkII homolog in Emericella
(Aspergillus) nidulans.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + protein -> ADP + O-phosphoprotein.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CAMK SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M74120; AAB97502.1; -.
DR EMBL; AF054580; AAD22581.1; -.
DR HSP; Q63450; 1A06.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
```

```
DR pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW Phosphorylation; ATP-binding.
FT DOMAIN 23 278 PROTEIN KINASE.
FT NP_BIND 29 37 ATP (BY SIMILARITY).
FT BINDING 50 50 ATP (BY SIMILARITY).
FT ACT_SITE 142 142 BY SIMILARITY.
FT DOMAIN 291 301 CALMODULIN-BINDING (BY SIMILARITY).
SQ SEQUENCE 414 AA; 46889 MW; 1E8D58A1C0B2F18C CRC64;

Query Match 25.9%; Score 651; DB 1; Length 414;
Best Local Similarity 38.3%; Pred. No. 6e-34;
Matches 149; Conservative 78; Mismatches 126; Indels 36; Gaps 11;

Qy 20 RKTFFIMEVLGSAFSEVFLVKQRTGKLFALKCIKKSPAFRDSLSLENEIAVLKKIKH 77
Dy 20 KSLYREGRTLGAAGTYG---IVREADSCSGKVAVKIILKRNVRGNRMVYDELDLQKLNH 76
Qy 78 ENIVLEDIYSTTHYLLVMQLVSGGELFDRILRGVYTEKDSALVIOQVLSAVKYLVH 137
Dy 77 PHIVHFVDWFESKDFYIVTQLATGGELFDRICEYKFTKDSQTIROVLDAVNYLHQ 136
Qy 138 GIVHRLDKPENLLYTPENSKIMTIDGLSKMEQN--GIMSTAGTGPYVAPEVLAQKP 195
Dy 137 NIVHRLDKPENLLYTPROLDLSQVLADFGIARMLDNPAEVLTSMAAGSFGYAAPEVLMK 196
Qy 196 YKAVDCWSIGVITYLLCGYPPFYEEETESKLFEEKIKEGYEFESPFWDIDISESAKDFIC 255
Dy 197 HGKAVDWSIGVITYLLCGYSPFRSENLTDLIECSRGRVVFHRYKWDVSKDAKDFIL 256
Qy 256 HLEKDPNRYTCERKALSHPWIDGNTALHRDIYPSVLSIQKNFAKSKWRQAFNAAVH 315
Dy 257 SLLQVDPAPQRTSEALKHPWLKGSASDRDLLP-----EIRAYIARSLKRGIEIKLAN 312
Qy 316 HMRKLMHNLHSPGVPRVENRPP--ETOASSTRSPSPETITEAPVL--DHSVALPAL 371
Dy 313 RTELAKMQ-----EDEEDIPSAVDVQASEASDKSG-----LSPFPALSTENHPAST 362
Qy 372 QLPQCHGRRTAPGG---RSLNCLVNGSL 397
Dy 363 ----GNGE-----SGGTKKRSLSKIARGAI 383

RESULT 8
KPSH_HUMAN STANDARD; PRT; 424 AA.
ID KPSH_HUMAN
AC P11801; Q9NY19;
DT 01-OCT-1989 (Rel. 12, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase H1 (EC 2.7.1.37) (PSK-H1).
GN PSKH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20541713; PubMed=11087665;
RA Brede G., Solheim J., Troen G., Prydz H.;
RT "Characterization of PSKH1, a novel human protein serine kinase with
centrosomal, golgi, and nuclear localization.";
RL Genomics 70:82-92(2000).
RN [2]
RP SEQUENCE OF 199-348 FROM N.A.
RX MEDLINE=87092414; PubMed=2948189;
RA Hanks S.K.;
RT "Homology probing: identification of cDNA clones encoding members of
```







DE Calcium/calmodulin-dependent protein kinase type II beta chain (EC  
DE 2.7.1.123) (CaM-kinase II beta chain) (CaM kinase II beta subunit)  
DE (CaM-II beta subunit).  
DE CAMK2B OR CAMKB OR CAM2.  
GN Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxId=9606;  
[1]  
RP SEQUENCE FROM N.A. (ISOFORMS 0; 1; 4; 6 AND 7).  
RP TISSUE=Brain;  
RC MEDLINE=20317016; PubMed=10858498;  
RX Wang P., Wu Y., Zhou T.H., Sun Y., Pei G.;  
RA "Identification of alternative splicing variants of the beta subunit  
RT of human Ca(2+)-calmodulin-dependent protein kinase II with different  
RT activities.";   
RL FEBS Lett. 475:107-110(2000).  
[2]  
RP SEQUENCE FROM N.A. (ISOFORM 3).  
RP TISSUE=Skeletal muscle;  
RA Leddy J.J., Salih M., Tuana B.S.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
[3]  
RP SEQUENCE FROM N.A. (ISOFORMS 0 AND 4).  
RP TISSUE=Brain;  
RA Li G.Y., Cooper N.G.F.;  
RT "Molecular cloning and sequencing of human calcium/calmodulin  
RT dependent protein kinase II beta subunit";  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
[4]  
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 4).  
RP TISSUE=Insulinoma;  
RX MEDLINE=20277365; PubMed=10819240;  
RA Rochlitz H., Voigt A., Lankat-Buttgereit B., Goke B., Heimberg H.,  
RA Nauck M.A., Schiemann U., Schatz H., Pfeiffer A.F.;  
RT "Cloning and quantitative determination of the human  
RT Ca2+/calmodulin-dependent protein kinase II (CaMK II) isoforms in  
RT human beta cells.";   
RL Diabetologia 43:465-473(2000).  
[5]  
RP SEQUENCE OF 302-603 FROM N.A. (ISOFORM 1).  
RP MEDLINE=97214619; PubMed=9060999;  
RA Tombes R.M., Krystal G.W.;  
RT "Identification of novel human tumor cell-specific CaMK-II variants.".  
RL Biochim. Biophys. Acta 1355:281-292(1997).  
CC -1- FUNCTION: THIS KINASE MAY PLAY A ROLE IN NEUROTRANSMISSION.  
CC -1- CATALYTIC ACTIVITY: ATP + protein = ADP + O-phosphoprotein.  
CC -1- ENZYME REGULATION: AUTOPHOSPHORYLATION OF CAM-KINASE II PLAYS AN  
CC IMPORTANT ROLE IN THE REGULATION OF THE KINASE ACTIVITY.  
CC -1- SUBUNIT: COMPOSED OF FOUR DIFFERENT CHAINS: ALPHA, BETA, GAMMA,  
CC AND DELTA.  
CC -1- ALTERNATIVE PRODUCTS: 7 ISOFORMS; 0/BETA1, 1/BETAL/BETA'E, 2/BETA2/  
CC 3 (SHOWN HERE), 4/BETA4/BETA'E, 6/BETA6 AND 7/BETA7; ARE PRODUCED  
CC BY ALTERNATIVE SPLICING.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC CAMK SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collabora-  
CC between the Swiss Institute of Bioinformatics and the EMBL outstat-  
CC the European Bioinformatics Institute. There are no restrictions on  
CC use by non-profit institutions as long as its content is in no  
CC modified and this statement is not removed. Usage by and for commer-  
CC entities requires a license agreement (See <http://www.isb-sib.ch/anno>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AF078803; AAD42035.1; -  
CC EMBL; AF081572; AAD42036.1; -  
CC EMBL; AF083419; AAD42038.1; -  
CC EMBL; AF081924; AAD42037.1; -  
CC EMBL; AF140350; AAD42070.1; -  
CC EMBL; U23460; AAC99802.1; -  
CC EMBL; AF112472; AAD03744.1; -  
CC EMBL; AF112471; AAD03743.1; -





QY 192 AORPYKAVDCWSIGVITYILLGYPFYETEESKLEKIKEGYEFESPFWDIDISESAK 251  
: || || | : || || || || || || || : : || : || || || : ||  
Db 186 RKDPYKGPVDLWACGVILYLLVGYPPFWDEDOHRLYQOIKAGAYDFPSPEWDTVTPEAK 245  
QY 252 DFICHLLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIOKNF-AKSKWRQAFNA 310  
|| : | : || : || : || || || : : : || : || : | : ||  
Db 246 DLINKMLTINPSKRITAAEALKHPWISHRSTVASCWHROETVDCCLKFNARRKLGAILT 305  
QY 311 AAVVHHM--RKLHMLHSPGVREVENRPPETQASETSRPSPEITITE 357  
: | : || : || : | : || : : || : || : || : ||  
Db 306 TMLATRNFGGKSGGNKNDGVKESSESTNTTIEDETKVRQEIIKYTE 355

Search completed: March 14, 2003, 14:08:58  
Job time : 17 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:07:31 ; Search time 18 Seconds  
(without alignments)  
2542.221 Million cell updates/sec

Title: US-09-960-643-2  
Perfect score: 2513  
Sequence: 1 MGRKEEDCSSWKKQTNR.....VKASGSSHCRCAGQGVCLIM 476

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1155.5	46.0	374	1 S50193	Ca2+/calmodulin-de
2	1152.5	45.9	370	1 S57347	Ca2+/calmodulin-de
3	1004	40.0	348	2 T37321	Ca2+/calmodulin-de
4	880	35.0	310	2 B88640	protein K07A9.2 [i
5	736	29.3	504	2 I56542	calmodulin-binding
6	703	28.0	421	2 T30814	calmodulin-binding
7	697.5	27.8	469	1 S17656	Ca2+/calmodulin-de
8	692.5	27.6	474	1 TVRTC4	Ca2+/calmodulin-de
9	692.5	27.5	502	2 I52637	Ca2+/calmodulin-de
10	691.5	27.5	301	1 A40811	myosin-light-chain
11	684	27.2	473	1 A50306	Ca2+/calmodulin-de
12	651	25.9	414	2 JN0323	Ca2+/calmodulin-de
13	630	25.1	335	2 T50290	calmodulin kinase
14	604.5	24.1	589	2 S68470	Ca2+/calmodulin-de
15	603	24.0	708	2 T23616	hypothetical prote
16	599.5	23.9	516	1 JU0270	Ca2+/calmodulin-de
17	599	23.8	478	1 A30355	Ca2+/calmodulin-de
18	599	23.8	530	2 D44412	Ca2+/calmodulin-de
19	598.5	23.8	509	2 B44412	calmodulin-depende
20	593	23.6	580	2 T40939	probable Ca-calmol
21	592	23.6	556	2 JC5636	Ca2+/calmodulin-de
22	591.5	23.5	533	1 A34366	Ca2+/calmodulin-de
23	588	23.4	447	2 B40896	Ca2+/calmodulin-de
24	588	23.4	478	1 S04365	Ca2+/calmodulin-de
25	583.5	23.2	547	2 T23614	hypothetical prote
26	582.5	23.2	518	1 S43845	Ca2+/calmodulin-de
27	581.5	23.1	518	1 B46619	Ca2+/calmodulin-de
28	579.5	23.1	542	1 A45025	Ca2+/calmodulin-de
29	578.5	23.0	542	1 A26464	Ca2+/calmodulin-de

30	577.5	23.0	446	2 A40896	Ca2+/calmodulin-de
31	573	22.8	554	2 T05476	calcium-dependent
32	569.5	22.7	504	2 T38226	probable serine-th
33	566.5	22.5	527	1 A31908	Ca2+/calmodulin-de
34	559.5	22.3	490	1 S71776	calcium-dependent
35	558	22.2	821	1 A39616	protein kinase RAD
36	556.5	22.1	610	1 A49082	calcium-dependent
37	553	22.0	490	2 T08873	calcium-dependent
38	546.5	21.7	520	2 F85059	probable calcium d
39	544.5	21.7	583	2 H84810	probable calcium d
40	544	21.6	443	2 T33690	hypothetical prote
41	543.5	21.6	319	2 I38138	protein-serine kin
42	543.5	21.6	508	1 A43713	calcium-dependent
43	543	21.6	495	1 S46284	calcium-dependent
44	541.5	21.5	553	1 T02139	calcium-dependent
45	541.5	21.5	591	2 S54788	calcium-stimulated

ALIGNMENTS

RESULT 1

S50193  
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) I - rat  
N: Alternate names: CaMKI  
C: Species: Rattus norvegicus (Norway rat)  
C: Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C: Accession: S50193; A49682; A46038  
R: Cho, F.S.; Phillips, K.S.; Bogucki, B.; Weaver, T.E.  
Biochim. Biophys. Acta 1224, 156-160, 1994  
A: Title: Characterization of a rat cDNA clone encoding calcium/calmodulin-dependent  
A: Reference number: S50193; MUID: 95035115; PMID: 7948038  
A: Accession: S50193  
A: Status: preliminary  
A: Molecule type: mRNA  
A: Residues: 1-374 <CHO>  
A: Cross-references: EMBL: L26288; NID: g439613; PIDN: AAA66944.1; PID: g439614  
R: Picciotto, M.R.; Czernik, A.J.; Nairn, A.C.  
J. Biol. Chem. 268, 26512-26521, 1993  
A: Title: Calcium/calmodulin-dependent protein kinase I. cDNA cloning and identifiat  
A: Reference number: A49682; MUID: 94075341; PMID: 8253780  
A: Accession: A49682  
A: Status: preliminary  
A: Molecule type: mRNA  
A: Residues: 1-111; G', 113-117, 'R', 119-308, 'R', 310-322, 'HOPG', 327, 'T', 329, 'TDS' <PIC>  
A: Cross-references: GB: I24907; NID: g406112; PIDN: AAA19670.1; PID: g406113  
R: Mochizuki, H.; Ito, T.; Hidaka, H.  
J. Biol. Chem. 268, 9143-9147, 1993  
A: Title: Purification and characterization of Ca2+/calmodulin-dependent protein kin  
A: Reference number: A46038; MUID: 93232082; PMID: 8386178  
A: Accession: A46038  
A: Status: preliminary  
A: Molecule type: protein  
A: Residues: 12-36, 'T' <MOC>  
A: Experimental source: cerebrium  
A: Note: sequence extracted from NCBI backbone (NCBIP: 129927)  
C: Superfamily: Ca2+/calmodulin-dependent protein kinase I; protein kinase homology  
C: Keywords: ATP; autophosphorylation; calmodulin binding; phosphoprotein; phosphotra  
F: 18-276/Domain: protein kinase homology <KIN>  
F: 26-34/Region: protein kinase ATP-binding motif  
F: 293-299/Region: autoinhibitory  
F: 302-314/Region: calmodulin binding  
F: 177/Binding site: phosphate (Thr) (covalent) (by autophosphorylation) #status pred

Query Match	46.08	Score 1155.5	DB 1	Length 374
Best Local Similarity	69.68	Pred. No. 6.9e-44		
Matches 222	Conservative 43	Mismatches 51	Indels 3	Gaps 3
QY	12	WKQTNTNRKTFIFWVLGSGAFSEVFLVKORLTGKLFALKCI-KKSPAFRDSLENEIA	70	
		: :		

Db 69 VLHKIKHPNIVALDDIYESGGHLYLMQLVSGGELFDRIVEKGFYTERDASRLIFQV LDA 128  
QY 131 VKYLHENGIVHRDLKPNENLYITPENSKIMITDFGLSKMEQNG-IMSTACGTPGYVAPE 189  
Db 129 VKYLDHGLTVHRDLKPNENLYISLDEDSKIMISDFGLSKMEDPGSVLSTACGTPGYVAPE 188  
QY 190 VLAQPKYSKAVDCWSIGVITYILLCGYPPFYETESKLFKEIKEGYEFESPFWDIDISES 249  
Db 189 VLAQPKYSKAVDCWSIGVITAYILLCGYPPFYENDAKLFEQILKAEYFEDSPYWDIDISDS 248  
QY 250 AKDFICHLLEKDPNERYTCEKALSHPWIDGNLTALHRIYPSVSLQIQKNFAKSKWRQAFN 309  
Db 249 AKDFIRHLEKDPKRFTEQALQHPWIAGDTALDKNITHOSVSEQIKKNFAKSKWKQAFN 308  
QY 310 AAQVHHMRKLNHLHSPG 328  
Db 309 ATAVVHRMRKQLGTSQEG 327

RESULT 2  
S57347  
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) I - human  
N:Alternate names: CamKI  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S57347  
R:Haribabu, B.; Hook, S.S.; Selbert, M.A.; Goldstein, E.G.; Tomhave, E.D.; Edelman, A.M.  
EMBO J. 14, 3679-3686, 1995  
A:Title: Human calcium-calmodulin dependent protein kinase I: cDNA cloning, domain struc  
A:Reference number: S57347; MUID:95369239; PMID:7641687  
A:Accession: S57347  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-370 <HAR>  
A:Cross-references: EMBL:L41816; NID:G790789; PIDN:AAA99458.1; PID:G790790  
C:Genetics:  
A:Gene: GDB:CAMK1  
C:Superfamily: Ca2+/calmodulin-dependent protein kinase I; protein kinase homology  
C:Keywords: ATP; autophosphorylation; calmodulin binding; phosphoprotein; phosphotransfe  
F:18-276/Domain: protein kinase homology <KIN>  
F:26-34/Region: protein kinase ATP-binding motif  
F:293-299/Region: autoinhibitory  
F:302-314/Region: calmodulin binding  
F:177/Binding site: phosphate (Thr) (covalent) (by autophosphorylation) #status predicted

Query Match 45.9%; Score 1152.5; DB 1; Length 370;  
Best Local Similarity 55.8%; Pred. No. 9.2e-44;  
Matches 239; Conservative 46; Mismatches 62; Indels 81; Gaps 9;

QY 12 WKQQTNIKRFTFMELVGLSGAFSEVFLVKQLRTGKLFALKCI-KKSPAFRDSLSLENEIA 70  
Db 10 W-KQADIRDIYDFRDVLGTGAFSEVLAEDRKTQKLVAKICAEALGEGSGMENEIA 68  
QY 71 VLKKIKHENIVTLEDIESTHYLYVMQLVSGGELFDRILRGVYTERDASLVIOQVLSA 130  
Db 69 VLHKIKHPNIVALDDIYESGGHLYLMQLVSGGELFDRIVEKGFYTERDASRLIFQV LDA 128  
QY 131 VKYLHENGIVHRDLKPNENLYITPENSKIMITDFGLSKMEQNG-IMSTACGTPGYVAPE 189  
Db 129 VKYLDHGLTVHRDLKPNENLYISLDEDSKIMISDFGLSKMEDPGSVLSTACGTPGYVAPE 188  
QY 190 VLAQPKYSKAVDCWSIGVITYILLCGYPPFYETESKLFKEIKEGYEFESPFWDIDISES 249  
Db 189 VLAQPKYSKAVDCWSIGVITAYILLCGYPPFYENDAKLFEQILKAEYFEDSPYWDIDISDS 248  
QY 250 AKDFICHLLEKDPNERYTCEKALSHPWIDGNLTALHRIYPSVSLQIQKNFAKSKWRQAFN 309  
Db 249 AKDFIRHLEKDPKRFTEQALQHPWIAGDTALDKNITHOSVSEQIKKNFAKSKWKQAFN 308  
QY 310 AAQVHHMRKLNHLHSPGVRPEVNPETQASSETSRSPPSEITITEAPVLDHVSALPA 369

Db 309 ATAVVHRMRKQLG-----TS----- 324  
QY 370 LTQLPCQHRRPTAPGGRSLNCLVNGSLHISSSLVPHHQGSLAAGPCGCCSSCLNIGSKG 429  
Db 325 -----QEGQQTASHGELLTPVAGGP-----AAG-C-CCRDC----- 354  
QY 430 KSSYCSEP 437  
Db 355 ---CVP 358  
RESULT 3  
T37321  
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) I - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jul-2000  
C:Accession: T37321  
R:Eto, K.; Takahashi, N.; Kimura, Y.; Masuho, Y.; Arai, K.; Muramatsu, M.; Tokumitsu, J. Biol. Chem. 274, 22556-22562, 1999  
A:Title: Ca2+/calmodulin-dependent protein kinase cascade in Caenorhabditis elegans.  
A:Reference number: Z21686; MUID:99357789; PMID:10428833  
A:Accession: T37321  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-348 <ETO>  
A:Cross-references: EMBL:AB021864; NID:G5672677; PIDN:BAA82674.1; PID:G5672678  
A:Experimental source: strain Bristol N2; embryonic stage  
C:Genetics:  
A:Note: cmk-1  
C:Superfamily: Ca2+/calmodulin-dependent protein kinase I; protein kinase homology  
C:Keywords: ATP; autophosphorylation; calmodulin binding; phosphoprotein; phosphotran  
Query Match 40.0%; Score 1004; DB 2; Length 348;  
Best Local Similarity 57.3%; Pred. No. 2.5e-37;  
Matches 185; Conservative 67; Mismatches 69; Indels 2; Gaps 2;  
QY 19 IRKTFIFMEVLGSGAFSEVFLVKQLR-TGKLFALKCI-KKSPAFRDSLSLENEIAVLKKIK 76  
Db 18 IREKYDFRDVLGTGAFSKVFLAESKDAGQYAVKCIDKKALKGKESLENEIKVLKRLR 77  
QY 77 HENIVTLEDIESTHYLYVMQLVSGGELFDRILRGVYTERDASLVIOQVLSAVKYLHE 136  
Db 78 HNNIVQLFDTDEKQFVFLVMELVTGGELFDRIVAKSYTEQDASNLIRQLVLEAVGFMD 137  
QY 137 NGIVHRDLKPNENLYITPENSKIMITDFGLSKMEQNGIMSTACGTPGYVAPEVLAQPKY 196  
Db 138 NGVHRDLKPNENLYITPENSKIMISDFGLSKTDSGVMTACGTPGYVAPEVLQPKY 197  
QY 197 SKAVDCWSIGVITYILLCGYPPFYETESKLFKEIKEGYEFESPFWDIDISESAKDFICH 256  
Db 198 GKAVDWSIGVITAYILLCGYPPFYDESANLFAQIKGEYFEDAPYWDQISGSAKDFITH 257  
QY 257 LLEKDPNERYTCEKALSHPWIDGNLTALHRIYPSVSLQIQKNFAKSKWRQAFNAAVVHH 316  
Db 258 LMCCDPEARFTCDALSHPWISGNTAYTHDIGHVAVHLKSLAKRNWKAYNAALAIQ 317  
QY 317 MRKLHMLHSPGVRPEVNPPE 339  
Db 318 LQMLRLSSNSNRLQKQASQOQPE 340

RESULT 4  
B88640  
protein K07A9.2 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-Aug-2001  
C:Accession: B88640  
R:anonymous, the C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_



RESULT 7  
S17656  
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) IV - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence.revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S17656; A29878; I49571  
R:Jones, D.A.; Glod, J.; Wilson-Shaw, D.; Hahn, W.E.; Sikela, J.M.  
FEBS Lett. 289, 105-109, 1991  
A:Title: cDNA sequence and differential expression of the mouse Ca(2+)/calmodulin-depend  
A:Reference number: S17656; MUID:91372388; PMID:1893997  
A:Accession: S17656  
A:Molecule type: mRNA  
A:Residues: 1-469 <JON>  
A:Cross-references: EMBL:X58995; NID:g50366; PIDN:CAAA1741.1; PID:g50367  
R:Sikela, J.M.; Hahn, W.E.  
Proc. Natl. Acad. Sci. U.S.A. 84, 3038-3042, 1987  
A:Title: Screening an expression library with a ligand probe: Isolation and sequence of  
A:Reference number: A29878; MUID:87204263; PMID:3033675  
A:Accession: A29878  
A:Molecule type: mRNA  
A:Residues: 315-469 <SIX>  
A:Cross-references: GB:M16206; NID:g200360; PIDN:AAA39933.1; PID:g387512  
R:Sikela, J.M.; Law, M.L.; Kao, F.  
Genomics 4, 21-27, 1989  
A:Title: Chromosomal localization of the human gene for brain Ca2+/calmodulin-dependent  
A:Reference number: I49571; MUID:89122027; PMID:2536634  
A:Accession: I49571  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 250-277, 'CFG', 281-301, 'T', 303-338, 'X', 340-469 <RES>  
A:Cross-references: GB:J03057; NID:g192366; PIDN:AAA37366.1; PID:g192367  
A:Experimental source: brain  
C:Superfamily: Ca2+/calmodulin-dependent protein kinase; protein kinase homology  
C:Keywords: alternative splicing; ATP; calmodulin binding; phosphotransferase; serine/th  
F:40-296/Domain: protein kinase homology <KIN>  
F:48-56/Region: protein kinase ATP-binding motif  
F:306-469/Product: calipermin #status predicted <CSP>  
F:318-337/Region: calmodulin binding #status predicted  
F:71/Active site: Lys #status predicted

Query Match	27.8%	Score 697.5	DB 1	Length 469
Best Local Similarity	42.7%	Pred. No. 7.2e-24		
Matches 152	Conservative 60	Mismatches 121	Indels 23	Gaps 7
Qy	9	CSMKKQTTNIRKTF-----IFMEV---LGSAFSEVFLVQRQLTGKLFALKCI	54	
Db	14	CSSTASTENLVDPYWDGNSNRDPLGDFFVESELGRGATSIYVRCQKGTQRPYALKVL	73	
Qy	55	KSPAFRDSLENEIAVLKTKKHENIVTLEDIYESTHYLVMLVSGGELFDRLRGV	114	
Db	74	KKT--VDKKIVRTETGIVLLRSLHNPNIKLKEIFETPTETISLVLELVTTGGELFDRIVEKGY	131	
Qy	115	YTEKDASLVIOOVLNAVYLVHENGIVHRDLKPNLLVLTPEENSKIMITDFGLSK-MEQN	173	
Db	132	YSERDARDVAQVQILEAVAYLVHENGIVHRDLKPNLLVLTAPDAPLKIADFGLSKIVEHQ	191	
Qy	174	GIMSTACGTCGYAPEVLAQXPYKAVDCWSIGVITVILLCGYPFPFYEEET-ESKLFPKIK	232	
Db	192	VLMTVCGTCGYCAPEILRGCAYPEVDMNSVGIIITVILLCGFEFFYDERGDQFMFRIL	251	
Qy	233	EGYVEFPFWDIDISAKDFICHLLEKDPNERYTCBKALSHPWIDGNTA--LHRDIYPS	290	
Db	252	NCEYFISFPMWDEVSLNAKDLVKLLVLDPPKRLTTFQALQHPWTKAANFVHMD---T	308	
Qy	291	VSLQIQNFASKVRQAFNAVAHVHMRKLMHNLHSPGVPEVNRPPETQASETS	346	
Db	309	AQKLLQEFNARRKLLKAAVKAVASSRLGSSASSSSTSIQENHKASSDPPTQDAKDS	364	

RESULT 8  
TVRUC4  
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) IV - rat

N:Alternate names: Ca2+/calmodulin-dependent protein kinase Gr  
N:Contains: calipermin  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Mar-1993 #sequence.revision 31-Mar-1993 #text\_change 11-Jun-1999  
C:Accession: A41103; A41237; A32865; A41250; A32035; A60255; I53706  
R:Ohmstede, C.A.; Bland, M.M.; Merrill, B.M.; Sahyoun, N.  
Proc. Natl. Acad. Sci. U.S.A. 88, 5784-5788, 1991  
A:Title: Relationship of genes encoding Ca(2+)/calmodulin-dependent protein kinase Gr  
A:Reference number: A41103; MUID:91288548; PMID:1648230  
A:Accession: A41103  
A:Molecule type: DNA  
A:Residues: 47-141, 'NE', 144-474 <OH1>  
A:Cross-references: GB:M74488; NID:g203219; PIDN:AAA40845.1; PID:g203220  
A:Note: this sequence has been revised in reference A41237  
A:Note: part of this sequence was confirmed by sequencing of cDNA to mRNA  
R:Ohmstede, C.A.; Bland, M.M.; Merrill, B.M.; Sahyoun, N.  
Proc. Natl. Acad. Sci. U.S.A. 88, 9375, 1991  
A:Reference number: A41237  
A:Accession: A41237  
A:Molecule type: DNA  
A:Residues: 142-143 <OH2>  
A:Cross-references: GB:M63334  
A:Note: this is a revision to the sequence from reference A41103  
R:Ohmstede, C.A.; Jensen, K.F.; Sahyoun, N.E.  
J. Biol. Chem. 264, 5866-5875, 1989  
A:Title: Ca(2+)/calmodulin-dependent protein kinase enriched in cerebellar granule ce  
A:Reference number: A32865; MUID:89174647; PMID:2538431  
A:Accession: A32865  
A:Molecule type: mRNA  
A:Residues: 250-474 <OH3>  
A:Cross-references: GB:J04600; NID:g206172; PIDN:AAA41867.1; PID:g206173  
R:Means, A.R.; Cruzalegui, F.; LeMagueresse, B.; Needleman, D.S.; Slaughter, G.R.; On  
Mol. Cell. Biol. 11, 3960-3971, 1991  
A:Title: A novel Ca(2+)/calmodulin-dependent protein kinase and a male germ cell-spec  
A:Reference number: A41250; MUID:91304387; PMID:1649385  
A:Accession: A41250  
A:Molecule type: mRNA  
A:Residues: 1-371, 'M', 373-408, 'Q', 410-474 <MEA>  
A:Cross-references: GB:M64757  
R:Ono, T.; Slaughter, G.R.; Cook, R.G.; Means, A.R.  
J. Biol. Chem. 264, 2081-2087, 1989  
A:Title: Molecular cloning sequence and distribution of rat calipermin, a high affini  
A:Reference number: A32035; MUID:89123272; PMID:2914893  
A:Accession: A32035  
A:Molecule type: mRNA  
A:Residues: 306-371, 'M', 373-474 <ONI>  
A:Cross-references: GB:J04446; NID:g203642; PIDN:AAA40990.1; PID:g203643  
R:Ono, T.; Means, A.R.  
Adv. Exp. Med. Biol. 255, 263-268, 1989  
A:Title: Calipermin is a testis specific calmodulin-binding protein closely related t  
A:Reference number: A60255; MUID:90144189; PMID:2618865  
A:Accession: A60255  
A:Molecule type: protein  
A:Residues: 335-363 <ON2>  
A:Note: the amino end of calipermin was blocked  
R:Bland, M.M.  
Gene 137, 351-352, 1993  
A:Title: Identification of alternate 5' untranslated regions in the gene encoding Ca2  
A:Reference number: I53706; MUID:94131312; PMID:8299971  
A:Accession: I53706  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-37 <RES>  
A:Cross-references: GB:LI6999; NID:g310086; PIDN:AAA17443.1; PID:g310087  
C:Comment: Ca2+/calmodulin-dependent protein kinase IV is enriched in cerebellar gran  
und in sperm cells.  
C:Superfamily: Ca2+/calmodulin-dependent protein kinase; protein kinase homology  
C:Keywords: alternative splicing; ATP; calmodulin binding; phosphotransferase; serine  
F:40-296/Domain: protein kinase homology <KIN>  
F:48-57/Region: protein kinase ATP-binding motif  
F:306-474/Product: calipermin #status predicted <CSP>  
F:318-337/Region: calmodulin binding #status predicted  
F:71/Active site: Lys #status predicted

```
Query Match 27.6%; Score 692.5; DB 1; Length 474;
Best Local Similarity 40.8%; Pred. No. 1.2e-23;
Matches 158; Conservative 62; Mismatches 120; Indels 47; Gaps 10;

QY 9 CSSWKKQTNIKRKF-----IPMEV---LGSGAFSEVFLVKORLTKGLFALKCI 54
DQ 14 CSSVTSSTENLVPDYWDIGSKRDPLSDFFEVESELGRGATSIYVRCKQKGTQKPYALKVL 73

QY 55 KKSAPFROSSLENEIAVLKKIKHENIVTLEDIYESTHYLYMOLVSGGELFDRILRGV 114
DQ 74 KKT--VDKKIVRTEIGVLRSHPNLIKKLKEIFETPTISLVLELTGTGELFDRIVERGY 131

QY 115 YTEKASLVIOQVLSAVKYLHENGIVHRDLKPENLLYLTPPENSKIMITDFGLSK-MEQN 173
DQ 132 YSERDAADAVKQILEAVYLHENGIVHRDLKPENLLYATPADPLKLTADFGLSKIVEHQ 191

QY 174 GIMSTACGTPGYVAVPEVLAQRPYSKAVDCWSIGVTIYLLCGYPFPYEET-ESKLFPEKIK 232
DQ 192 VLMKTVCGTPGYCAPEILRGCAIGPEVDMWSVGIIITYILLCGFFPFYDERGQDFMFRIL 251

QY 233 EGYEYFESFPWDDISESAKDFICHLEKDPNERVTCERKALSHPWIDGNTA--LHRDIYPS 290
DQ 252 NCEYFISFWDEVSINAKDLVKLLVDPKRLTTFQALQHPWWTGKAANFVHMD---T 308

QY 291 VSLQIQKNFAKSKWROAFNA-----AAVVHMRKLLHMLNLSHSPGVRP-----EV 333
DQ 309 AQKKLQEFNARRKLAAKAVVASSRLGSSASSHTNIQESNKASSEAPQADGKDKTDPL 368

QY 334 ENR-----PPETOASSETSRSPSP 353
DQ 369 ENKIQAQDHEAAKAAADETMKLOSEEV 395

RESULT 9
I52637
Ca2+/calmodulin-dependent protein kinase IV beta polypeptide - rat
C:Species: Rattus sp. (rat)
C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 18-Jun-1999
C:Accession: I52637
R:Sakagami, H.; Kondo, H.
Brain Res. Mol. Brain Res. 19, 215-218, 1993
A:Title: Cloning and sequencing of a gene encoding the beta polypeptide of Ca2+/calmodulin-
A:Reference number: I52637; MUID:94018484; PMID:8412563
A:Accession: I52637
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-502 <RES>
A:Cross-references: GB:S65840; NID:g425383; PIDN:AAB28372.1; PID:g425384
C:Superfamily: Ca2+/calmodulin-dependent protein kinase; protein kinase homology
C:Keywords: ATP; calmodulin binding
F:68-324/Domain: protein kinase homology <KIN>
F:76-84/Region: protein kinase ATP-binding motif

Query Match 27.6%; Score 692.5; DB 2; Length 502;
Best Local Similarity 40.8%; Pred. No. 1.3e-23;
Matches 158; Conservative 62; Mismatches 120; Indels 47; Gaps 10;

QY 9 CSSWKKQTNIKRKF-----IPMEV---LGSGAFSEVFLVKORLTKGLFALKCI 54
DQ 42 CSSVTSSTENLVPDYWDIGSKRDPLSDFFEVESELGRGATSIYVRCKQKGTQKPYALKVL 101

QY 55 KKSAPFROSSLENEIAVLKKIKHENIVTLEDIYESTHYLYMOLVSGGELFDRILRGV 114
DQ 102 KKT--VDKKIVRTEIGVLRSHPNLIKKLKEIFETPTISLVLELTGTGELFDRIVERGY 159

QY 115 YTEKASLVIOQVLSAVKYLHENGIVHRDLKPENLLYLTPPENSKIMITDFGLSK-MEQN 173
DQ 160 YSERDAADAVKQILEAVYLHENGIVHRDLKPENLLYATPADPLKLTADFGLSKIVEHQ 219

QY 174 GIMSTACGTPGYVAVPEVLAQRPYSKAVDCWSIGVTIYLLCGYPFPYEET-ESKLFPEKIK 232
DQ 220 VLMKTVCGTPGYCAPEILRGCAIGPEVDMWSVGIIITYILLCGFFPFYDERGQDFMFRIL 279
```

```
QY 233 EGYEYFESFPWDDISESAKDFICHLEKDPNERVTCERKALSHPWIDGNTA--LHRDIYPS 290
DQ 280 NCEYFISFWDEVSINAKDLVKLLVDPKRLTTFQALQHPWWTGKAANFVHMD---T 336

QY 291 VSLQIQKNFAKSKWROAFNA-----AAVVHMRKLLHMLNLSHSPGVRP-----EV 333
DQ 337 AQKKLQEFNARRKLAAKAVVASSRLGSSASSHTNIQESNKASSEAPQADGKDKTDPL 396

QY 334 ENR-----PPETOASSETSRSPSP 353
DQ 397 ENKMQAQDHEAAKAAADETMKLOSEEV 423

RESULT 10
A40811
myosin-light-chain kinase (EC 2.7.1.117) A - slime mold (Dictyostelium discoideum)
N:Alternate names: MLCK-A
C:Species: Dictyostelium discoideum
C>Date: 10-Apr-1992 #sequence_revision 21-Jan-1997 #text_change 11-Jun-1999
C:Accession: A40811; A37125
R:Tan, J.L.; Spudich, J.A.
J. Biol. Chem. 266, 16044-16049, 1991
A:Title: Characterization and bacterial expression of the Dictyostelium myosin light
A:Reference number: A40811; MUID:91340753; PMID:1651931
A:Accession: A40811
A:Molecule type: mRNA
A:Residues: 1-301 <TAN>
A:Cross-references: GB:M64176; NID:g1498249; PIDN:AAB06337.1; PID:g1498250
R:Tan, J.L.; Spudich, J.A.
J. Biol. Chem. 265, 13818-13824, 1990
A:Title: Dictyostelium myosin light chain kinase. Purification and characterization.
A:Reference number: A37125; MUID:90337997; PMID:2380188
A:Accession: A37125
A:Molecule type: protein
A:Residues: 9-12,'1',14-19;163-167,'s',169-179;192-198 <TA2>
C:Genetics:
A:Gene: mlka
C:Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-thre
d in itself
A:Pathway: cytokinesis; fruiting body formation
A:Note: not activated by Ca2+/calmodulin in contrast with MLCK from higher eukaryote.
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; autophosphorylation; cell division; phosphoprotein; phosphotransfer
F:6-265/Domain: protein kinase homology <KIN>
F:14-22/Region: protein kinase ATP-binding motif
F:267-295/Domain: inhibitory #status predicted <INH>
F:37/Active site: Lys #status predicted
F:296/Binding site: phosphate (Thr) (covalent) (by autophosphorylation) #status predi

Query Match 27.5%; Score 691.5; DB 1; Length 301;
Best Local Similarity 50.2%; Pred. No. 8.7e-24;
Matches 135; Conservative 52; Mismatches 79; Indels 3; Gaps 2;

QY 17 TNIRKTFEFMEVLGSGAFSEVFLVKORLTGKLFALCKKSPAFRD--SSLENEIAVLKK 74
DQ 2 TEVEKIYEFKEELGRGAFSIYLGCKNQTKQRYAIVKINKSELGKDYKXNLMKEVDILKK 61

QY 75 IKHENIVTLEDIYESTHYLYMOLVSGGELFDRILRGVYTEKASLVIOQVLSAVKYL 134
DQ 62 VNHFNIIALKELFDTPEKLYLMVELVTGGEFLDKIVEKGSYSEADAANLVKKIVSAVGYL 121

QY 135 HENGIVHRDLKPENLLYLTPPENSKIMITDFGLSK-MEQNGIMSTACGTPGYVAVPEVLAQ 193
DQ 122 HGLNIVHRDLKPENLLKSKENHLEVAIADFGLSKIIGTGLVMTQACGTPSYVAPEVLNA 191

QY 194 KPSKAVDCWSIGVTIYLLCGYPFPYEETESKLFKEIKEGYEFESFPWDDISESAKDF 253
DQ 182 TGYDKVEYDMSIGVTIYLLCGFPFPFYGDIVPEIFEQIMEVNFEPEEYWGGLSKEAKDF 241

QY 254 ICHLEKDPNERVTCERKALSHPWIDGNTA 282
DQ 254 ICHLEKDPNERVTCERKALSHPWIDGNTA 282
```





Search completed: March 14, 2003, 14:10:05  
Job time : 21 secs



```
QY 293 TCAAGCATGAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCAACCCACTACTACC 352
|||||
Db 241 TCAAGCATGAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCAACCCACTACTACC 300
QY 353 TGGTCATGAGCTTGTCTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGCT 412
|||||
Db 301 TGGTCATGAGCTTGTCTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGCT 360
QY 413 ACACAGAGAAGATGCCAGTCTGGTGATCCAGAGGCTCTTCTGGCAGTGAATACCTAC 472
|||||
Db 361 ACACAGAGAAGATGCCAGTCTGGTGATCCAGAGGCTCTTCTGGCAGTGAATACCTAC 420
QY 473 ATGAGAATGGCATCTGCACAGAGACTTAAAGCCCGGAAAACTCTTTACCTTACCCCTG 532
|||||
Db 421 ATGAGAATGGCATCTGCACAGAGACTTAAAGCCCGGAAAACTCTTTACCTTACCCCTG 480
QY 533 AAGAGAACTCTAAGATCATGATGACTTGTGCTCTCAAGATGGAACAGAAATGGCA 592
|||||
Db 481 AAGAGAACTCTAAGATCATGATGACTTGTGCTCTCAAGATGGAACAGAAATGGCA 540
QY 593 TCATGTCCACTCCCTGTGGGACCCAGGCTACGTGGCTCCAGAGAGTGTGCGCCAGAAAC 652
|||||
Db 541 TCATGTCCACTCCCTGTGGGACCCAGGCTACGTGGCTCCAGAGAGTGTGCGCCAGAAAC 600
QY 653 CCTACAGCAAGGCTGTGATGCTGGTCCATCGGCGTCATCACCTACATATTTGCTGTG 712
|||||
Db 601 CCTACAGCAAGGCTGTGATGCTGGTCCATCGGCGTCATCACCTACATATTTGCTGTG 660
QY 713 GATACCCCGCTTCTATGAGAAAGCGAGTCTAAGCTTTTCGAGAGATCAAGGAGGCT 772
|||||
Db 661 GATACCCCGCTTCTATGAGAAAGCGAGTCTAAGCTTTTCGAGAGATCAAGGAGGCT 720
QY 773 ACTATGAGTTTGAGTCTCCATCTCGGATGACATTTCTGAGTCAGCCCAAGGACTTTATTT 832
|||||
Db 721 ACTATGAGTTTGAGTCTCCATCTCGGATGACATTTCTGAGTCAGCCCAAGGACTTTATTT 780
QY 833 GCCATTGCTTGAGAAGGATCGGAACGAGCGGTACACCTGTGAGAAGGCCCTTGAGTCATC 892
|||||
Db 781 GCCATTGCTTGAGAAGGATCGGAACGAGCGGTACACCTGTGAGAAGGCCCTTGAGTCATC 840
QY 893 CTGGATTGACGGAACACGGCCCTCCACCGGACATCTACCCATCAGTCAGCCTCCAGA 952
|||||
Db 841 CTGGATTGACGGAACACGGCCCTCCACCGGACATCTACCCATCAGTCAGCCTCCAGA 900
QY 953 TCCAGAAGAACTTTGCTAAGAGCAAGTGGAGCAAGCTTCAACGCGACAGCTGTGTGTC 1012
|||||
Db 901 TCCAGAAGAACTTTGCTAAGAGCAAGTGGAGCAAGCTTCAACGCGACAGCTGTGTGTC 960
QY 1013 ACCATATGAGGAAGCTACACATGAACCTGCACAGCCCGGGCGTCCGCCAGAGGTGGAGA 1072
|||||
Db 961 ACCATATGAGGAAGCTACACATGAACCTGCACAGCCCGGGCGTCCGCCAGAGGTGGAGA 1020
QY 1073 ACAGGCGGCTGAAACTCAAGCTCAGAAACCTCTAGACCCAGCTCCCTCAGATCAACA 1132
|||||
Db 1021 ACAGGCGGCTGAAACTCAAGCTCAGAAACCTCTAGACCCAGCTCCCTCAGATCAACA 1080
QY 1133 TCACGAGGACGCTGTCTGTGACCAACAGTGTAGCACTCCCTGCCCTGACCCAAATACCCCT 1192
|||||
Db 1081 TCACGAGGACGCTGTCTGTGACCAACAGTGTAGCACTCCCTGCCCTGACCCAAATACCCCT 1140
QY 1193 GCCAGATGGCGCGGCGGCTACCTGCCCCTGGTGGCAGGTCCCTCAACTGCTGTGTCATG 1252
|||||
Db 1141 GCCAGATGGCGCGGCGGCTACCTGCCCCTGGTGGCAGGTCCCTCAACTGCTGTGTCATG 1200
QY 1253 GCTCCCTCCACATCAGCAGACGCTGTGGTCCCATCATCAGGGTCCCTGCGCGCGGGG 1312
|||||
Db 1201 GCTCCCTCCACATCAGCAGACGCTGTGGTCCCATCATCAGGGTCCCTGCGCGCGGGG 1260
QY 1313 CCTGTGGCTGTGCTCCAGCTGCTGAACATTTGGAGCAAAAGGAAAGTCTCTACTGCT 1372
|||||
Db 1261 CCTGTGGCTGTGCTCCAGCTGCTGAACATTTGGAGCAAAAGGAAAGTCTCTACTGCT 1320
```

```
QY 1373 CTGAGCCACACTCTCTCAAAAAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTCATGG 1432
|||||
Db 1321 CTGAGCCACACTCTCTCAAAAAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTCATGG 1380
QY 1433 TACAGTTAAAGCCAGTGGCAGCTCCCACTGCCGCGCAGGCGAGCTGGAGTCTGTCTCA 1492
|||||
Db 1381 TACAGTTAAAGCCAGTGGCAGCTCCCACTGCCGCGCAGGCGAGCTGGAGTCTGTCTCA 1440
QY 1493 TTATGTGATTCCTGGAGCCTGTGCCTATGTCTACTGTCAATTTTTCAGGAGACATATTCAACT 1552
|||||
Db 1441 TTATGTGATTCCTGGAGCCTGTGCCTATGTCTACTGTCAATTTTTCAGGAGACATATTCAACT 1500
QY 1553 CCTCTGCTCTTCCAAACCTGTGTCTATCCGGCAGAGGAGGAGGAGCAAGTGGAG 1612
|||||
Db 1501 CCTCTGCTCTTCCAAACCTGTGTCTATCCGGCAGAGGAGGAGGAGCAAGTGGAG 1560
QY 1613 CAGGCTTACGAGGAGCAGTTTCTGGCCAGAAGCACCAAGCTCTGCCAGCGGGGAGCC 1672
|||||
Db 1561 CAGGCTTACGAGGAGCAGTTTCTGGCCAGAAGCACCAAGCTCTGCCAGCGGGGAGCC 1620
QY 1673 CCTCATAGGAGGCCAGGAGGAGGCCCAAGGCGTTAGAAGCCTTGTGAAGCTGTGAGCA 1732
|||||
Db 1621 CCTCATAGGAGGCCAGGAGGAGGCCCAAGGCGTTAGAAGCCTTGTGAAGCTGTGAGCA 1680
QY 1733 GGAGAAGCGGTGCCACCAAGCTTCCAGGTCTCCCTGACCTGCTCTATGCCCCAC 1790
|||||
Db 1681 GGAGAAGCGGTGCCACCAAGCTTCCAGGTCTCCCTGACCTGCTCTATGCCCCAC 1738
```

## RESULT 2

US-09-935-464-2

; Sequence 2, Application US/09935464

; Publication No. US20030027153A1

; GENERAL INFORMATION:

; APPLICANT: Meyer, Joanne

; APPLICANT: Barrington-Martin, Rory

; APPLICANT: Parker, Alexander

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCH

; FILE REFERENCE: 3322/1H702 US1

; CURRENT APPLICATION NUMBER: US/09/935,464

; CURRENT FILING DATE: 2001-08-23

; PRIOR APPLICATION NUMBER: US 09/757,300

; NUMBER OF SEQ ID NOS: 90

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 1383

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-935-464-2

Query Match 54.7%; Score 1339.6; DB 9; Length 1383;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1342; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 70 ATGGGTGCAAGGAAGAAGATGATGTCAGTTCCTGTGGAAGAAACAGACCACACATCCCGG 129

Db 1 ATGGGTGCAAGGAAGAAGATGATGTCAGTTCCTGTGGAAGAAACAGACCACACATCCCGG 60

QY 130 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTTTTCCTGGTG 189

Db 61 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTTTTCCTGGTG 120

QY 190 AAGCAAGACTGACTGGGAAAGCTTTTGTCTCTGAAGTGCATCAAGAGTCACCTGCTTC 249

Db 121 AAGCAAGACTGACTGGGAAAGCTTTTGTCTCTGAAGTGCATCAAGAGTCACCTGCTTC 180

QY 250 CGGACAGACGCTGGAGAATGAGATTGCTGTGTTGAAAAGATCAAGCATGAAACATT 309

Db 181 CGGACAGACGCTGGAGAATGAGATTGCTGTGTTGAAAAGATCAAGCATGAAACATT 240

QY 310 GTGACCCCTGGAGGACATCTATGAGACCAACCCACTACTACCTGGTCTGACGCTGTT 369



Qy	2170	TCCAAAGTGGAAATAGAAAGATTTCATGAGTAAAGGCTCGAAGGAATTCCTTATCTGGCC	2229
Db	148131	TCCAAAGTGGAAATAGAAAGATTTCATGAGTAAAGGCTCGAAGGAATTCCTTATCTGGCC	148190
Qy	2230	ACATGTCTCCGTGCACACACCCAAATGGAGTTAAACCTTGGAAAGTTGACATATTTTAAATGTC	2289
Db	148191	ACATGTCTCCGTGCACACACCCAAATGGAGTTAAACCTTGGAAAGTTGACATATTTTAAATGTC	148250
Qy	2290	TGCGAGGAGTTCTAAATCTCGCCTCTGTTCCCTTTCTCTCTCTGAAAGTCCAGCACACCA	2349
Db	148251	TGCGAGGAGTTCTAAATCTCGCCTCTGTTCCCTTTCTCTCTCTGAAAGTCCAGCACACCA	148310
Qy	2350	TTCTTGTGCTTCCCGCAGTTTCCCTCGCCCTCCACCCCTCAGCTTCATGCTCAGTGTGTGG	2409
Db	148311	TTCTTGTGCTTCCCGCAGTTTCCCTCGCCCTCCACCCCTCAGCTTCATGCTCAGTGTGTGG	148370
Qy	2410	CTTAATAAAATGGACATATTTTTTCTCTA	2437
Db	148371	CTTAATAAAATGGACATATTTTTTCTCTA	148398

```

RESULT 4
US-10-024-036B-1
; Sequence 1, Application US/10024036B
; Publication No. US20030028004A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 68730 and 69112, Prote
; TITLE OF INVENTION: Molecules and Uses Th
; FILE REFERENCE: MPI2000-521PR(M)
; CURRENT APPLICATION NUMBER: US/10/024,036B
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/256222
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1772

```

```

: LENGTH: 1772
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: 5'UTR
: LOCATION: (1)...(168)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (169)...(1242)
: FEATURE:
: NAME/KEY: 3'UTR
: LOCATION: (1243)...(1772)
: FEATURE:
: NAME/KEY: unsure
: LOCATION: 1
: OTHER INFORMATION: n may be A, T, G, or C
: FEATURE:
: NAME/KEY: unsure
: LOCATION: 3
: OTHER INFORMATION: n may be A, T, G, or C
: FEATURE:
: NAME/KEY: unsure
: LOCATION: 5
: OTHER INFORMATION: n may be A, T, G, or C
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (6)...(6)
: OTHER INFORMATION: n may be A, T, G, or C
: US-10-024-0368-1

```

Query Match	20.6%	Score 504.6;	DB 9;	Length 1772;
Best Local Similarity	71.3%;	Pred. No. 8.5e-134;		
Matches 695;	Conservative 0;	Mismatches 274;	Indels 6;	Gaps 2;
Qv	66	GGCAATGGGTGCGAAGGGAAGATGACTGCAGTTCTCTGGGAAGAACACGACCACCAACAT	125	

[illegible]

RESULT 5  
US-10-024-036B-3  
; Sequence 3, Application US/10024036B  
; Publication No. US20030028004A1  
; GENERAL INFORMATION:

```
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 68730 and 69112, Protein Kinase
; FILE REFERENCE: MPI2000-521PIR(M)
; CURRENT APPLICATION NUMBER: US/10/024,036B
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/258222
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1074
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-024-036B-3

Query Match      20.5%; Score 502.2; DB 9; Length 1074;
Best Local Similarity 71.3%; Pred. No. 3e-133;
Matches 692; Conservative 0; Mismatches 273; Indels 6; Gaps 2;

QY 70 ATGGTCGAAGGAAGAAGATGACTGCAGTTCTCTGGAAGAAACAGACACACATCCGG 129
Db 1 ATGGCCGGGAGAACGGCAGAGCAGCTCTCTGGAAGAAAGCAAGCTGAAGACATCAAG 60

QY 130 AAAACCTTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAACTTTTCTCGTGGTG 189
Db 61 AAGATCTTCAGTTCAAAGAGACCCCTCGGAACGGGGCCCTTTCCGAAGTGGTTTAGCT 120

QY 190 AAGCAAGACTGACTGGGAAGCTTTTGTGCTGTAAGTGATC---AAGAAGTCACTGCC 246
Db 121 GAAGAGAAGGCAACTGCAAGCTCTTTGCTGTGAAGTGATCCCTAAGAAGGCGCTGAAG 180

QY 247 TTCGGGACAGCAGCTGGAGATGAGATTGCTGTGTTTGAAGAAAGATCAAGCATGAAC 306
Db 181 GGCAGGAAGCAGCATAGAGATGAGATGAGCGCTCTGAGAAAGATTGAAGCATGAAAT 240

QY 307 ATTGTGACCTGGAGACATCTATGAGAGCACCACCTACTACTCTACCTGTGTCATGCAGCTT 366
Db 241 ATGTTGCCCTGGAAGACATTTATGAAGCCCAATCACTGTACTTGTGTCATGCAGCTG 300

QY 367 GTTCTGTGGGGAGCTCTTTGACCGGATCTGTGGAGCGGGGTGCTPACACAGAGAAGAT 426
Db 301 GTCTCGGGTGGAGAGCTGTTTGACCGGATAGTGGAGAAGGGGTTTATACAGAGAAGAT 360

QY 427 GCCAGTCTGTGATCAGCAGGCTTTGTCGGCAGTCAATACCTACATCAGATGAGCATC 486
Db 361 GCCAGCACTGTATCCGCAAGTCTTTGAGCGCGGTGTACTATCTCCACAGAAATGGGCATC 420

QY 487 GTCCACAGAGACTTAAAGCCCGAAAAACCTGTCTTACCTTTACCCCTGAAGAGAACTCTAAG 546
Db 421 GTCCACAGAGACTCAAGCCCGGAAATCTTGTACTACAGTCAAGATGAGGAGTCCAA 480

QY 547 ATCATGATCACTGACCTTTGGTCTGTCCAGATGGAACAGAATGG---CATCATGTCCACT 603
Db 481 ATAAATGATCAGTCACTTTGGATTGTCAAAAATGGAGGGCAAAAGGAGATGTGATGTCCACT 540

QY 604 GCCTGTGGACCCAGGCTACGTGGCTCCAGAAAGTCTGGCCCAAGAAACCCCTACAGCAAG 663
Db 541 GCCTGTGGAACTCCAGGCTATGTGCTCTGTAAGTCTCTCGCCAGAAACCTTACAGCAA 600

QY 664 GCTGTGGATTGCTGGTCCATCGGCGTCATCACTACATATTGCTGTGGATACCCCGG 723
Db 601 GCGGTGACTGTGTTGCTTCCAGGAGTGATGCGTACATCTTGTCTGCGGCTACCCCTCT 660

QY 724 TTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTTACTAGAGTTT 783
Db 661 TTTTATGATGAAAATGACTCCAAGCTCTTTTGAGCAGATCCCTCAAGGCGGAATATGAGTT 720

QY 784 GAGTCTCCATTCTGGGATGACATTTCTGAGTCAGCCAAAGCACTTTATTTGCCACTTGCTT 843
Db 721 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAAGACTTCATTTCGGAACCTGATG 780

QY 844 GAGAAGGATCCGAACAGCGGTACACCTCTGTGAGAGGGCCTTGTAGTCAATCCCTGGATTGAC 903
```

```
Db 781 GAGAAGGACCCGAATAAAGAATACACGTGTGAGCAGGCGAGCTCGGCACCCATGATCGCT 840
QY 904 GGAACACAGGGCCCTCCACCGGACATCTACCATCAGTCAGCTCAGCTCCAGATCCAGAAGAAC 963
Db 841 GGTGACACAGCCCTCAACAAAAACATCCACGAGTCCGTTCAGCGCCAGATCCGGAAGAAC 900
QY 964 TTTGCTAAGACGAAGTGGAGGCAAGCCCTTCAACGAGCAGCAGCTGTGGTGCACCATGAGG 1023
Db 901 TTTGCCAAGACAAATGGAGACAAGCATTTAATGCGACGCGCGTCTGTGAGACATATGAGA 960
QY 1024 AAGCTACACAT 1034
Db 961 AACTACACT 971

RESULT 6
US-09-935-464-7
; Sequence 7, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSY
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 501
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-7
```

```
Query Match      20.4%; Score 499.4; DB 9; Length 501;
Best Local Similarity 99.8%; Pred. No. 1.2e-132;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1525 CTGCAATTTTCAGGAGACATATTCAACTCTCTCTCTTCCAAACCTGGTCTATCCGG 1584
Db 1 CTGCAATTTTCAGGAGACATATTCAACTCTCTCTCTTCCAAACCTGGTCTATCCGG 60

QY 1585 CAGAGGAGGAAGGCAGAGCAAGTGGAGCGGCTTAGCAGGACAGTTTCTGGCCAGAA 1644
Db 61 CAGAGGAGGAAGGCAGAGCAAGTGGAGCGGCTTAGCAGGACAGTTTCTGGCCAGAA 120

QY 1645 GCACCAGCTCTGCCAGCGGGCAGCCCTCATAGAGGGCCAGAGGGAGCCCCAAGG 1704
Db 121 GCACCAGCTCTGCCAGCGGGCAGCCCTCATAGAGGGCCAGAGGGAGCCCCAAGG 180

QY 1705 CGTAGAACCTTGTGAACTGTGAGCAGGAGAGACGGTGCACCAGCTTCCAGGTCTC 1764
Db 181 CGTAGAACCTTGTGAAAGCTGTGAGCAGGAGAGCGGTGCCACCAGCTTCCAGGTCTC 240

QY 1765 CTTGACCTGCTGCTTCTATGCCCCACACCTACGTGCGCTGGCTGTGTGCAGTGTAGTA 1824
Db 241 CTTGACCTGCTGCTTCTATGCCCCACACCTACGTGCGCTGGCTGTGTGCAGTGTAGTA 300

QY 1825 GATAGCTCTCGCTGGGTCTGTGCTGTTGTCGTAAGGCTTAATGGCTGGCCAGGCT 1884
Db 301 GATAGCTCTCGCTGGGTCTGTGCTGTTGTCGTAAGGCTTAATGGCTGGCCAGGCT 360

QY 1885 GTGTCACCTTCTCCAAAGCAAAAGCCATATGGAGCATCTACCCAGACTGCCACTCTGCACAC 1944
Db 361 GTGTCACCTTCTCCAAAGCAAAAGCCATATGGAGCATCTACCCAGACTGCCACTCTGCACAC 420

QY 1945 ACTCACTCCCACTCTCAAGCTTCCAACTCTTGSCCAGATTGGCTCATTAATGTCGTT 2004
```

Db 421 ACTACTCCACCTCTCAAGCCTCAACCTCTTGCCAGATTGGGCTCATTAATGTCGTT 480

QY 2005 GCCTGCCCATCTGCATGAATG 2025  
|||||

Db 481 GCCTGCCCATCTGCATGAATG 501

RESULT 7

US-09-835-788A-6

; Sequence 6, Application US/09835788A

; Patent No. US20020077458A1

; GENERAL INFORMATION:

; APPLICANT: Ni et al.

; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides, a

; FILE REFERENCE: PT018P1

; CURRENT APPLICATION NUMBER: US/09/835,788A

; PRIOR FILING DATE: 2001-04-17

; PRIOR APPLICATION NUMBER: PCR/US00/28666

; PRIOR FILING DATE: 2000-10-17

; PRIOR APPLICATION NUMBER: 60/159,585

; PRIOR FILING DATE: 1999-10-18

; PRIOR APPLICATION NUMBER: 60/167,246

; PRIOR FILING DATE: 1999-11-24

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 6

; LENGTH: 1578

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-835-788A-6

Query Match 20.3%; Score 497; DB 10; Length 1578;

Best Local Similarity 71.6%; Pred. No. 1.2e-131;

Matches 682; Conservative 0; Mismatches 265; Indels 6; Gaps 2;

QY 88 GATGACTCGAGTCTCTGGGAAGAACAGACACCACCAATCCGGAACCTTCATTTTATG 147  
|||

Db 15 GAGAGCAGCTCTCTCGGAAAAGCAAGCTGAAGACATCAAGAAGATCTTCGAGTTCAA 74  
|||

QY 148 GAAGTCTGGGATCAGGAGCTTTCTCAGAAATTTTCTCGGTGAAGCAAGACTGACTGGG 207  
|||

Db 75 GAGACCTTCGGAACCGGGSCCTTTCCGAAGTGGTTTAGCTGAAGAGAGCAACTGSC 134  
|||

QY 208 AAGCTCTTTGCTCTGAAGTGCATC---AAGAAGTCACTGCCCTCCGGAGCAGAGCCTG 264  
|||||

Db 135 AAGCTCTTTGCTCTGAAGTGTATCCCTAAGAAGGCGCTGAAGGGCAAGGAAGCAGCATA 194  
|||||

QY 265 GAGATGAGATTGCTGTGTTGAAAAGATCAAGCATGAAACATTTGACCTCGAGGAGAC 324  
|||||

Db 195 GAGATGAGATAGCCGTCTCGAAGAAGATTAAGCATGAAATATTTGTCCTGGAGAC 254  
|||||

QY 325 ATCTATGAGAGCACCACCACTACTACCTGGTCAATGCAGCTTTGTTCTGGTGGGAGCTC 384  
|||||

Db 255 ATTTAGAAAGCCCAATCACCTGTACTTTGGTCATGCGAGCTGGTGCCTGGAGAGCTG 314  
|||||

QY 385 TTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAGGATGCCAGTCTGGTATCCAG 444  
|||||

Db 315 TTTGACCGGATGATGAGAGGGGTTTTATACAGAGAAGGATGCCAGCACTCTGATCCGC 374  
|||||

QY 445 CAGGTCCTTCGCGAGTGAATACCTACATGAGATGCAATGCAATGCAAGCAAGCACTTAAG 504  
|||||

Db 375 CAAAGTCTTGAGCGCGGTGACTATCTCCACAGATGGGATCGTCCACAGAGACCTCAAG 434  
|||||

QY 505 CCCGAAACCTCTTTACCTTTACCCCTGAGAGAACTCTAAGATCATGATCACTGACITTT 564  
|||||

Db 435 CCCGAAATCTCTTGTAATGATGAGGAGTCCAAATATATGATCACTGACITTT 494  
|||||

QY 565 GGTCTGTCCAAGATGGAACAGAAATGG---CATCATGTCACCTGTCCTGTGGAGCCCGG 621  
|||||

Db 495 GGATTTGTCAAAATGGAGGGCAAGAGATGTGATGTCCACTGCTGCGCACTCCAGG 554  
|||||

QY 622 TACGTGGCTCCAGAAATGCTGTGCCCCAGAAACCTTACAGAAAGGCTGTGGTGTGCTG 681

Db 555 TATGTCGCTCCTGAAGTCTCGCCAGAAACCTTTACAGCAAGCCGTTGACTGCTGGTCC 614  
|||||

QY 682 ATCGGCGCTCATACATATTGCTCTGTGGATACCCCGCTTCTATGAAGAAACGGAG 741  
|||||

Db 615 ATCGGAGTGATGCTACATCTGCTCTGGGTACCCCTCTTTTATGATGAATGAC 674  
|||||

QY 742 TCTAAGCTTTTCGAGAAGATCAAGGAGGGCTACTATGAGTTTGAGTCTCACTCTGGGAT 801  
|||||

Db 675 TCCAAGCTCTTTGAGCAGATCCTCAAGGCGGAATATGAGTTTGACTCTCCCTACTGGGAT 734  
|||||

QY 802 GACATTTCTGAGTCAGCAAGGACTTTATTTGCCACTTGTTCGCAAGGATCCGACGAG 861  
|||||

Db 735 GACATCTCCGACTCTGAAAAGACTTCATTCCGAACCTGATGGAGAAGACCCGGAATAA 794  
|||||

QY 862 CGGTACACCTGTGAGAAAGCCTTGAGTCACTCCCTCGGATTCACGGAACACAGCGCCTCCAC 921  
|||||

Db 795 AGATACAGCTGTGAGCAGGACGCTCGGCCCATGGATCGCTGGTACACAGCCCTCAAC 854  
|||||

QY 922 CGGACATCTACCCATCAGTCAGCTCCAGATCCAGAACACTTTGCTAAGACGAAGTGG 981  
|||||

Db 855 AAAAACATCCACGAGTCCGTGAGCGCCAGATCCGGAACAACTTTGCAAGAGCAATGG 914  
|||||

QY 982 AGCAAGCCTTCACGACGAGCAGCTGTGTCACCATGAGGAGCTACACAT 1034  
|||||

Db 915 AGCAAGCATTTAATGCCACGCGCTGTGAGACATATGAGAAACTACACCT 967  
|||||

RESULT 8

US-09-935-464-46

; Sequence 46, Application US/09935464

; Publication No. US20030027153A1

; GENERAL INFORMATION:

; APPLICANT: Meyer, Joanne

; APPLICANT: Barrington-Martin, Rory

; APPLICANT: Parker, Alexander

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYC

; FILE REFERENCE: 3322/1H702 US1

; CURRENT APPLICATION NUMBER: US/09/935,464

; CURRENT FILING DATE: 2001-08-23

; PRIOR APPLICATION NUMBER: US 09/757,300

; PRIOR FILING DATE: 2001-01-09

; NUMBER OF SEQ ID NOS: 90

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 46

; LENGTH: 480

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; OTHER INFORMATION: n=a or g or c or t/u

US-09-935-464-46

Query Match 19.6%; Score 479; DB 9; Length 480;

Best Local Similarity 99.8%; Pred. No. 8e-127;

Matches 479; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 41 CATCCTCAGAAAGCTTCAACTCTGGAGCAATGGGTGCGAAAAGAAAGATGACTGCAAGTT 100  
|||||

Db 1 CATCCTCAGAAAGCTTCAACTCTGGAGCAATGGGTGCGAAAAGAAAGATGACTGCAAGTT 60  
|||||

QY 101 CCTGGAAGAAACAGACACCACCAATCCGGAACCTTCATTTTATGGAAGTCTGGGAT 160  
|||||

Db 61 CCTGGAAGAAACAGACACCACCAATCCGGAACCTTCATTTTATGGAAGTCTGGGAT 120  
|||||

QY 161 CAGGAGCTTTCTCAGAAAGTTTCTGCTGAGCAAGAAAGACTGACTGGGAAGCTCTTTGCTC 220  
|||||

Db 121 CAGGAGCTTTCTCAGAAAGTTTCTGCTGAGCAAGAAAGACTGACTGGGAAGCTCTTTGCTC 180  
|||||

QY 221 TGAAGTGCATCAAGAAGTCACTGCCCTTCGCGGACAGCAGCCTGGAGAATGAGATTGCTG 280  
|||||

Db 181 TGAAGTGCATCAAGAAGTCACTGCCCTTCGCGGACAGCAGCCTGGAGAATGAGATTGCTG 240  
|||||

QY 281 TGTGAAAGATCAAGCATGAACATCTGTGACCTGGAGGACATCTATGAGACACCA 340  
|||||  
Db 241 TGTGAAAGATCAAGCATGAACATCTGTGACCTGGAGGACATCTATGAGACACCA 300  
|||||  
QY 341 CCCACTACTAGCTGGTCATGAGCTTTCTGTGGGGAGCTTTTGACCGGATCCTGG 400  
|||||  
Db 301 CCCACTACTAGCTGGTCATGAGCTTTCTGTGGGGAGCTTTTGACCGGATCCTGG 360  
|||||  
QY 401 AGCGGGTGTCTACACAGAAGATGCCAGTCTGTGTATCCAGCAGTCTTGTGCGCAG 460  
|||||  
Db 361 AGCGGGTGTCTACACAGAAGATGCCAGTCTGTGTATCCAGCAGTCTTGTGCGCAG 420  
|||||  
QY 461 TGAATACCTACATGAGATGGCATGCTCCACAGAGACTTTAAAGCCCGAAACCTGCTT 520  
|||||  
Db 421 TGAATACCTACATGAGATGGCATGCTCCACAGAGACTTTAAAGCCCGAAACCTGCTT 480  
|||||  
RESULT 9  
US-09-935-464-47  
; Sequence 47, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300  
; PRIOR FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 47  
; LENGTH: 467  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: n=a or g or c or t/u  
US-09-935-464-47  
Query Match 17.7%; Score 434.2; DB 9; Length 467;  
Best Local Similarity 96.4%; Pred. No. 5.1e-114;  
Matches 450; Conservative 0; Mismatches 16; Indels 1; Gaps 1;  
QY 41 CATCCTCAGAGCTTCAACTCTGGAGCAATGGTGGAAAGGAGAGATGACTGCAGTT 100  
|||||  
Db 1 CATCCTCAGAGCTTCAACTCTGGAGCAATGGTGGAAAGGAGAGATGACTGCAGTT 60  
|||||  
QY 101 CCTGGAAGAAACAGACACCAACATCCGGAACCTTCATTTTATGGAAGTCTGGGAT 160  
|||||  
Db 61 CCTGGAAGAAACAGACACCAACATCCGGAACCTTCATTTTATGGAAGTCTGGGAT 120  
|||||  
QY 161 CAGAGCTTCTCAGAAGTTTCTGTGTGAAGCAAGACTGACTGGAGCTTTTGCTC 220  
|||||  
Db 121 CAGAGCTTCTCAGAAAGTTTCTGTGTGAAGCAAGACTGACTGGAGCTTTTGCTC 180  
|||||  
QY 221 TGAAGTGCATCAAGAAGTCACTCCCTCCGGGACAGCAGCCCTGGAGATGAGATGCTG 280  
|||||  
Db 181 TGAAGTGCATCAAGAAGTCACTCCCTCCGGGACAGCAGCCCTGGAGATGAGATGCTG 240  
|||||  
QY 281 TGTGAAAGATCAAGCATGAACATCTGTGACCTGGAGGACATCTATGAGACACCA 340  
|||||  
Db 241 TGTGAAAGATCAAGCATGAACATCTGTGACCTGGAGGACATCTATGAGACACCA 300  
|||||  
QY 341 CCCACTACTAGCTGGTCATGAGCTTTCTGTGGGGAGCTTTTGACCGGATCCTGG 399  
|||||  
Db 301 CCCACTACTAGCTGGTCATGAGCTTTCTGTGGGGAGCTTTTGACCGGATCCTGG 360  
|||||  
QY 400 GAGCGGGTGTCTACACAGAAGATGCCAGTCTGTGTATCCAGCAGGCTTTGTGCGCA 459  
|||||

Db 361 GAGCGGGTGTCTACACAGAAGATGCCAGNCTGGGTGATCCACANGTCTTGTGCGCA 420  
|||||  
QY 460 GTGAAATACCTACATGAGAATGGCATGCTCCACAGAGACTTTAAAGCC 506  
|||||  
Db 421 GTGAAATACCTACATGAGAATGGCATGCTCCACAGAGACTTTAAAGCC 467  
|||||  
RESULT 10  
US-09-817-181-1  
; Sequence 1, Application US/09817181  
; Patent No. US20020142427A1  
; GENERAL INFORMATION:  
; APPLICANT: MERKULOV, Gennady et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CL001189  
; CURRENT APPLICATION NUMBER: US/09/817,181  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1372  
; TYPE: DNA  
; ORGANISM: Human  
US-09-817-181-1  
Query Match 17.7%; Score 432.4; DB 10; Length 1372;  
Best Local Similarity 73.9%; Pred. No. 3.2e-113;  
Matches 563; Conservative 0; Mismatches 196; Indels 3; Gaps 1;  
QY 277 GCTGTGTTGAAAAGATCAAGCATGAAACATTTGTACCTCGGAGACATCTATGAGAGC 336  
|||||  
Db 204 GATGTTCTGGGCACGATCAAGCACCCCAACATTTGTAGCCCTGGATGACATCTATGAGAGT 263  
|||||  
QY 337 ACCAGCCACTACTACTGGTCATGAGCTTTGTTCTGTGGGGAGCTTTTGACCGGATC 396  
|||||  
Db 264 GGGGGCCACCTCTACCTCATGCGCTGTGTGGTGGGGAGCTTTTGACCGGATC 323  
|||||  
QY 397 CTGGAGCGGGTGTCTACACAGAAGATGCCAGTCTGTGTGATCCAGCAGCTTTGTGCG 456  
|||||  
Db 324 GTGGAAGAAAGCTTCTACACGAGCGGAGCGCCAGCCCTCATCTTCCAGGTGCTGAT 383  
|||||  
QY 457 GCAGTGAATACCTACATGAGAAATGGCATGCTCCACAGAGACTTTAAAGCCCGAAACCTG 516  
|||||  
Db 384 GCTGTGAAATACCTGACCTGAGCTGGCATTTGTACACCGGGATCTCAAGCCAGAGAAATCTG 443  
|||||  
QY 517 CTTTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCAGCTTTGGTCTGTCCAG 576  
|||||  
Db 444 CTGTACTACAGCTGGATGAAGACTCCAAATCATGATCTCCGACTTTGGGCTCTTCCAG 503  
|||||  
QY 577 ATGGAACAGAAATGGCA---TCATGTCCACTGCTGTGGGACCCAGGCTAGTGTGCTCCA 633  
|||||  
Db 504 ATGGAAGACCCGGGAGTGTGCTCTCCAGCCCTGTGGAACCTCCGGGATAGTGTGCTCCA 563  
|||||  
QY 634 GAAGTGTGCGCCCAAGAACCCCTACAGAGAGCTGTGGAATGCTGTGCTCATCGGCGTCA 693  
|||||  
Db 564 GAAGTGTGCGCCCAAGAACCCCTACAGAGAGCTGTGGAATGCTGTGCTCATAGTGTCA 623  
|||||  
QY 694 ACCTACATATGCTGTGTGATACCCCGCTTCTATGAAGAACGAGCTCTAAGCTTTTC 753  
|||||  
Db 624 GCCTACATCTTGTCTGTGGGTTACCTCTCTTCTATGAGAGATGATGCGCAACTCTTT 683  
|||||  
QY 754 GAGAAGATCAAGAGGCTTACTATGAGTTTGTGATCTCCATCTCGGATGACATTTCTGAG 813  
|||||  
Db 684 GAACAGATTTTGAAGCCGAGTACAGTTTGACTCTCTCTTACTGGAGACATCTCTGAC 743  
|||||  
QY 814 TCAGCCAGAGACTTTTATTTGCCATTTGCTGTGAGAGAGATCCGAGAGCGGTACACCTGT 873  
|||||  
Db 744 TCTGCCAAGATTTCTATCCGGCATTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 803  
|||||  
QY 874 GAGAAGGCTTTGAGTTCATCCCTGGATTGACGGAACACGCGCTTCCACCGGAGATCTAC 933  
|||||

Db 804 GAGCAGCCCTTGACGACCCATGGATTGCAGGAGATACAGCTCTAGATAAGAAATATCCAC 863  
QY 934 CCATCAGTCAGCCTCCAGATCCAGAAAGAACTTTGCTTAAGCAAGTGGAGGCAAGCCTTC 993  
Db 864 CAGTCGGTGAGTCAGCAGATCAAGAAAGAACTTTGCCAAGCAAGTGGAGCAAGCCTTC 923  
QY 994 AAGCAGCAGCTGTGTGCACCAATGAGAGAACTACATG 1035  
Db 924 AATGCCACGGCTGTGTGGCGCACATGAGGAAACTGCAGCTG 965

RESULT 11  
US-09-935-464-48  
; Sequence 48, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 48  
; LENGTH: 470  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: n-a or g or c or t/u  
US-09-935-464-48

Query Match 16.9%; Score 412.6; DB 9; Length 470;  
Best Local Similarity 96.2%; Pred. No. 7.7e-108;  
Matches 452; Conservative 0; Mismatches 14; Indels 4; Gaps 3;  
QY 41 CATCCTCAGAGCTTCAACTCTGGAGCAATGGTGCAGGAAAGAAAGATGACTGCAGTT 100  
Db 1 CATCCTCAGAGCTTCAACTCTGGAGCAATGGTGCAGGAAAGAAAGATGACTGCAGTT 60  
QY 101 CCTGGAAGAAACAGACACCAACAT--CGGAAACCTTCAATTTTATGGAAGTGTCTGGGA 159  
Db 61 CCTGGAAGAAACAGACACCAACATCCCGAAACCTTCAATTTTATGGAAGTGTCTGGGA 120  
QY 160 TCAGGAGCTTTCTCAGAA--GTTTTCTGTGTAAGCAAAAGTACTGCTGGGAAGCTCTTTG 217  
Db 121 TCAGGAGCTTTCTCAGAAAGTTTTCCCTGTTGTAAGCAAAAGTACTGCTGGGAAGCTCTTTG 180  
QY 218 CTCTGAAGTGCATCAAGAAATGCACTGCTTCCGGGACAGCCTGGAGAAATGAGATTG 277  
Db 181 CTCTGAAGTGCATCAAGAAATGCACTGCTTCCGGGACAGCCTGGAGAAATGAGATTG 240  
QY 278 CTGTGTTGAAAAGATCAAGCATGAAACATTTGTACCCCTGGAGGACATCTATGAGAGCA 337  
Db 241 CTGTGTTGAAAAGATCAAGCATGAAACATTTGTACCCCTGGAGGACATCTATGAGAGCA 300  
QY 338 CCACCCTACTACTACCTGGTTCATGCAAGTGGTCTGGTGGGAGCTCTTTGACCGG 393  
Db 301 CCACCCTACTACTACCTGGTTCATGCAAGTGGTCTGGTGGGAGCTCTTTGACCGG 356  
QY 396 CCACCCTACTACTACCTGGTTCATGCAAGTGGTCTGGTGGGAGCTCTTTGACCGGATC 396  
Db 301 CCACCCTACTACTACCTGGTTCATGCAAGTGGTCTGGTGGGAGCTCTTTGACCGGATC 360  
QY 397 CTGGAGCGGGTCTTACACAGAGAGGATGCCAGTCTGGTATCCAGCAGCTCTTGCG 456  
Db 361 CTGGAGCGGGTCTTACACAGAGAGGATGCCAGTCTGGTATCCAGCAGCTCTTGTCN 420  
QY 457 GCAGTGAATACCTACATGAGATGCAATGCTGCCACAGAGACTTAAAGCC 506  
Db 421 GCAGTGAATACCTACATGAGATGCAATGCTGCCACAGAGACTTAAAGCC 470

RESULT 12  
US-09-935-464-49  
; Sequence 49, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCH  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 49  
; LENGTH: 356  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: n-a or g or c or t/u  
US-09-935-464-49

Query Match 13.2%; Score 324.2; DB 9; Length 356;  
Best Local Similarity 98.0%; Pred. No. 1.2e-82;  
Matches 349; Conservative 0; Mismatches 4; Indels 3; Gaps 2;  
QY 41 CATCCTCAGAGCTTCAACTCTGGAGCAATGGTGCAGGAAAGAAAGATGACTGCAGTT 100  
Db 1 CATCCTCAGAGCTTCAACTCTGGAGCAATGGTGCAGGAAAGAAAGATGACTGCAGTT 60  
QY 101 CCTGGAAGAAACAGACACCAACAT--CGGAAACCTTCAATTTTATGGAAGTGTCTGGGA 159  
Db 61 CCTGGAAGAAACAGACACCAACATCCCGAAACCTTCAATTTTATGGAAGTGTCTGGGA 120  
QY 160 TCAGGAGCTTTCTCAGAA--GTTTTCTGTGTAAGCAAAAGTACTGCTGGGAAGCTCTTTG 217  
Db 121 TCAGGAGCTTTCTCAGAAAGTTTTCCCTGTTGTAAGCAAAAGTACTGCTGGGAAGCTCTTTG 180  
QY 218 CTCTGAAGTGCATCAAGAAATGCACTGCTTCCGGGACAGCCTGGAGAAATGAGATTG 277  
Db 181 CTCTGAAGTGCATCAAGAAATGCACTGCTTCCGGGACAGCCTGGAGAAATGAGATTG 240  
QY 278 CTGTGTTGAAAAGATCAAGCATGAAACATTTGTACCCCTGGAGGACATCTATGAGAGCA 337  
Db 241 CTGTGTTGAAAAGATCAAGCATGAAACATTTGTACCCCTGGAGGACATCTATGAGAGCA 300  
QY 338 CCACCCTACTACTACCTGGTTCATGCAAGTGGTCTGGTGGGAGCTCTTTGACCGG 393  
Db 301 CCACCCTACTACTACCTGGTTCATGCAAGTGGTCTGGTGGGAGCTCTTTGACCGG 356

RESULT 13  
US-09-935-464-50  
; Sequence 50, Application US/09935464  
; Publication No. US20030027153A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyer, Joanne  
; APPLICANT: Barrington-Martin, Rory  
; APPLICANT: Parker, Alexander  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCH  
; FILE REFERENCE: 3322/1H702 US1  
; CURRENT APPLICATION NUMBER: US/09/935,464  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: US 09/757,300  
; NUMBER OF SEQ ID NOS: 90

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-50

Query Match      13.0%; Score 319; DB 9; Length 319;
Best Local Similarity 100.0%; Pred. No. 3.6e-81;
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 CATCTCAGAGCTCAACTCTGGAGCAATGGGTGCGAAGCAAGAGATCACTGCAGTT 100
Db 1 CATCTCAGAGCTCAACTCTGGAGCAATGGGTGCGAAGCAAGAGATCACTGCAGTT 60
QY 101 CCTGGAAGAAACAGACACCAACATCCCGGAAAACCTTCATTTTATGGAAGTCTGGGAT 160
Db 61 CTGGAGNAACAGACACCAACATCCCGGAAAACCTTCATTTTATGGAAGTCTGGGAT 120
QY 161 CAGGAGCTTTCAGAAAGTTTCTCGTGAAGCAAAAGACTGACTGGGAAGCTCTTTGGTC 220
Db 121 CAGGAGCTTTCAGAAAGTTTCTCGTGAAGCAAAAGACTGACTGGGAAGCTCTTTGGTC 180
QY 221 TGAAGTCATCAAGAGTCACCTGCTTCCGGGACAGCAGCTGGAGAATGAGATTGCTG 280
Db 181 TGAAGTCATCAAGAGTCACCTGCTTCCGGGACAGCAGCTGGAGAATGAGATTGCTG 240
QY 281 TGTGAAAAGATCAAGCATGAAAACATTTGTGACCTCGGAGACATCTATGAGAGCACCA 340
Db 241 TGTGAAAAGATCAAGCATGAAAACATTTGTGACCTCGGAGACATCTATGAGAGCACCA 300
QY 341 CCCACTACTACCTGGTCAT 359
Db 301 CCCACTACTACCTGGTCAT 319

RESULT 14
US-09-764-868-189
; Sequence 189, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 189
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (59)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-189

Query Match      11.9%; Score 291; DB 9; Length 476;
Best Local Similarity 97.8%; Pred. No. 4.7e-73;
Matches 316; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY 1 TGGAGTGGGAGCTCAAGCAGGATCTTCCCGAGTCCTCGGATCTCAGAACCTTCAACT 60
Db 145 TGGAGTGGGAGCTCAAGCAGGATCTTCCCGAGTCCTCGGATCTCAGAACCTTCAACT 204
QY 61 CTGAGGCAATGGTCAAGGAAAGATGATGCTGAGTTCTCTGGAAGAAACAGACCACC 120
Db 205 CTGAGGCAATGGTCAAGGAAAGATGATGCTGAGTTCTCTGGAAGAAACAGACCACC 264
QY 121 AACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAGTT 180
Db 121 AACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAGTT 180
```

```
Db 265 AACATCGGAAAAACCTTCATTTTATGGAAGTCTGGATCAGGAGCTTCTCAGAAGTT 324
QY 181 TTCCTGTTGAAGCAAAAGACTGACTGGAGAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA 240
Db 325 TTCCTGTTGAAGCAAAAGACTGACTGGAGAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA 384
QY 241 CCTGCGCTTCGGGACAGCAGCCCTGGAGATGAGA-TTGCTGTGTTGAAAAAG-ATCAAGC 298
Db 385 CCTGCGCTTCGGGACAGCAGCCCTGGAGATGAGANTTGTCTGTGTTGAAAAAGTTTCAAGC 444
QY 299 ATGAAAACATTTGTGACCTGGAG 321
Db 445 ATGAAAACATTTGTGACCTGGG 467

RESULT 15
US-09-796-692-7777/c
; Sequence 7777, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THE
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7777
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (10)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (19)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (20)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-7777

Query Match      10.9%; Score 266.2; DB 9; Length 474;
Best Local Similarity 73.8%; Pred. No. 5.8e-66;
Matches 350; Conservative 0; Mismatches 121; Indels 3; Gaps 1;
```

Qy	410	TCATACAGAGAGGATGCCAGTCTGGTGATCCAGCAGGCTTGTGCGCAGTGAAATACC	469
Db	474	TCATACGGAGCGGAGCGCGCCCTCATCTTCAGGTGCTGGATGCTGTGAAATACC	415
Qy	470	TACATGAGAAATGCGATCGTCCACAGAGACTTAAAGCCCGAAAACCTGTCTTACCTTACCC	529
Db	414	TGCATGACCTGGGCATTTGATACACCGGATCTCAAGCCAGAGAATCTGCTGTACTACAGCC	355
Qy	530	CTGAAGAGAACTCTAAGATCATGATCAGTCTTGGTCTGTCCAGATGGAAACAAGTG	589
Db	354	TGGATGAAGACTCCAAAATCATGATCCGACTTTGGCCTCTCCAAGATGGAGGACCCGG	295
Qy	590	GCA---TCATGTCCACTGTGGGACCCAGGCTACGTGGCTCCAGAAGTGCCTGGGCC	646
Db	294	GCAGTGTGCTCTCCACCGCTGTGGAACTCCGGGATACGTGGCCCTGAAGTCTCTGGGCC	235
Qy	647	AGAAACCTTACAGCAAGGCTGTGGATTGCTGTCCATCGCGCTCATCACTACATATTGC	706
Db	234	AGAAGCCCTACAGCAAGGCTGTGGATTGCTGTCCATAGTGTGTCATCGCTACATCTTGC	175
Qy	707	TCGTGGATACCCCGTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGG	766
Db	174	TCTGCGGTTACCTCCCTTCTATGACGAGAATGATGCCAAACTCTTTGAACAGATTTTGA	115
Qy	767	AGGCTACTATAGTTTGTGATCTCCATCTGGGATGACATTTCTGAGTCAGCCAAGGACT	826
Db	114	AGCCGAGTACGAGTTTGACTCTCCTTACTGGGACGACATCTCTGACTGCGCAAGATT	55
Qy	827	TTATTGCCACTTGTGTGAGAGGATCCGAACGAGCGGTACACCTGTGAGAAGG	880
Db	54	TCATCCGGCACTTGATGGAGAGGACCCAGAGAAANNNGATTTCACNGTGAGCAGG	1

Search completed: March 14, 2003, 17:30:10  
Job time : 415 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 14:03:50 ; Search time 95 Seconds  
(without alignments)  
7899.348 Million cell updates/sec

Title: US-09-960-643-1  
Perfect score: 2447  
Sequence: 1 tggagtggaagctcaagcag.....ttttctctaaaaa 2447

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA.\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	418	17.1	1282	2	US-08-878-989-12
2	418	17.1	1282	4	US-09-272-796-12
3	337.2	13.8	362	4	US-08-905-223-89
4	243.6	10.0	3124	4	US-09-734-030-1
5	197.2	8.1	3471	2	US-08-715-568A-2
6	168.6	6.9	2061	4	US-09-800-960-1
7	163.4	6.7	2132	2	US-09-159-385-3
8	163.4	6.7	2132	4	US-09-186-277-3
9	145.8	6.0	1429	2	US-09-159-385-4
10	145.8	6.0	1429	4	US-09-186-277-4
11	135.6	5.5	1349	1	US-07-951-715A-20
12	135.6	5.5	1349	2	US-08-459-448A-20
13	135.6	5.5	1349	3	US-08-459-595A-20
14	135.6	5.5	1349	3	US-08-459-504B-20
15	135.6	5.5	1349	3	US-08-459-444-20
16	135.6	5.5	1349	4	US-09-547-422-20
17	123	5.0	1400	1	US-08-464-164-1
18	123	5.0	1400	1	US-08-338-057-1
19	123	5.0	1400	2	US-08-668-416-1
20	122.4	5.0	4935	2	US-08-631-097-3
21	122.4	5.0	5886	4	US-08-810-712-9
22	122.2	5.0	2637	4	US-09-735-934A-1
23	120.4	4.9	1735	4	US-09-746-694-3
24	120.4	4.9	1858	4	US-09-529-093A-1
25	119	4.9	1776	3	US-08-655-352-10
26	119	4.9	1776	4	US-09-258-016-10
27	119	4.9	1776	4	US-09-257-825B-10

28 114.8 4.7 1333 4 US-09-142-551A-1  
29 113.4 4.6 2514 3 US-08-655-352-1  
30 113.4 4.6 2514 4 US-09-258-016-1  
31 113.4 4.6 2514 4 US-09-257-825B-1  
32 108.4 4.4 4162 2 US-08-459-448A-26  
33 108.4 4.4 4162 3 US-08-459-595A-26  
34 108.4 4.4 4162 3 US-08-459-504B-26  
35 108.4 4.4 4162 3 US-08-459-444-26  
36 108.4 4.4 4162 4 US-09-547-422-26  
37 108.4 4.4 4165 1 US-07-951-715A-26  
38 107.2 4.4 2549 4 US-09-467-082-3  
39 106.8 4.4 2374 4 US-09-347-801-3  
40 106 4.3 8906 2 US-08-826-267-1  
41 104.6 4.3 2908 4 US-09-930-181-1  
42 101.8 4.2 2610 2 US-09-212-771-1  
43 101.8 4.2 2610 3 US-09-091-058-1  
44 99.6 4.1 1599 3 US-09-256-465-1  
45 99.6 4.1 1599 4 US-09-167-322-3

#### ALIGNMENTS

RESULT 1  
US-08-878-989-12  
; Sequence 12, Application US/08878989  
; Patent No. 5885803  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl G.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/878,989  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0321 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1282 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PROSNOT06

; CLONE: 827431  
US-08-878-989-12

Query Match 17.1%; Score 418; DB 2; Length 1282;  
Best Local Similarity 65.5%; Pred. No. 2.8e-109;  
Matches 626; Conservative 0; Mismatches 326; Indels 3; Gaps 1;

QY 105 GAAGAAACAGACACCAACATCCGGAAACCTTCATTTTATGGAAGTGTGGGATCAGG 164  
DB 212 GAAGAAACACACGGAGACATCAGCAGCGTCTACGAGATCCGCGAGAGGCTCGGCTCGG 271  
QY 165 ACCTTTCTCAGAAGTTTCCCTGCTGAGCAAGAGTACTGCTGGGAGCTCTTCTCTGAA 224  
DB 272 TGCCTTCTCCGAGGTGGCTGGCCAGAGCGGGGCTCCGACACCTCTGGGCCCTCAA 331  
QY 225 GTGCAT---CAAGAAGTCCACCTGCTCCGGACAGCAGCTCGAGAAATGAGATTGCTGT 281  
DB 332 GTGCATCCCAAGAAGGCCCTCCGGGCAAGGAGGCCCTGTGTGGAACGAGATCGCAGT 391  
QY 282 GTTGAAGAAGATCAAGCATGAAACATTTGTGACCCCTGGAGGACATCTATGAGACACCA 341  
DB 392 GCTCGTAGGATCAGTACCCCAACATCGTCGCTCGGAGGATGTCCACGAGAGCCCTTC 451  
QY 342 CCACCTACTACCTGGTGCAGCTGTGTTCTGTTGGGAGCTCTTTGACCGGATCCTTGA 401  
DB 452 CCACCTACTACCTGGCCAGACTGGTGACGGGTGGCGAGCTGTTGACCGCATCATGGA 511  
QY 402 GCGGGGTGTCTACACAGAGAAGATGCCAGTCTGGTGTATCCAGCAGCTCTTTGCGGCG 461  
DB 512 GCGGGGTCTTACACAGAGAAGATGCCAGCTCTGGTGTATCCAGCAGCTCTTTGCGGCG 571  
QY 462 GAAATACCTACATGATGATGTCATCGTCCACAGAGACTTAAAGCCGAAACCTGCTTTA 521  
DB 572 CTCTACTCTGACAGCCTGGGGATCTGTGACCGGGAGCTCAAGCCCGAAACCTCTCTGA 631  
QY 522 CTTTACCCTGAAGAGAACTTAAGATCATGATCACTGACTTTGGTGTGTCGAAGATGA 581  
DB 632 TGCCACGCGCTTTGAGGACTCGAAGATCATGTTCTGACTTTGGACTCTCCAAATCCA 691  
QY 582 ACAGATGGCATCATGCTCACTGCTGCTGGGACCCAGGCTACGTGGCTCCAGAAGTGT 641  
DB 692 GGCTGGGAACATGCTAGGACCGCTGTGGACCCCTTGAGATGTGGCCCGCAGAGCTCT 751  
QY 642 GCGCCAGAAACCTTACAGAAAGCTGTGGATGCTGTGCTTCCATCGGCTCATCACTACAT 701  
DB 752 GGAGCAGAAACCTTACGGAAGGCGGTAGATGTGTGGCCCTGGCGCTCATCTCCTACAT 811  
QY 702 ATTGCTCTGTGGATACCCCGCTTCTATGAGAAGCGAGTCTAAGCTTTTCGAGAGAT 761  
DB 812 CTTGCTGTGGGTACCCCGCTTCTACGAGAGGAGCCCTGAGCTCTTCAGCCAGAT 871  
QY 762 CAAGAGGGCTACTATGAGTTTGTAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGCAA 821  
DB 872 CTTGAGGCGCAGCTATGATTTGACTTTCTTCTGGGATGACATCTCAGAAATCAGCAA 931  
QY 822 GACCTTTATTTGCCATTTGCTTTGAGAAGATTCGAACGAGCGGTACACCTGTGAGAAGGC 881  
DB 932 AGACTTTATTTGGCACCTTCTGGAGCGAGACCTTCAAGAAGAGGTTTCACTGCCAACAGGC 991  
QY 882 CTTGAGTATCCTCTGGATTGACGAACACAGCGCCCTCCACGGGACATCTACCCATCAGT 941  
DB 992 CTTGGGAGCCTTTTGGATCTTTTGGGACACAGGCTTTGGCAGGAGCATCTTAGGGTTGT 1051  
QY 942 CAGCCTCCAGATCCAGAGAACTTTGCTTAAGAGCAAGTGGAGGCAAGCCTTCAACGAGC 1001  
DB 1052 CAGTGAGCAGATCCGAAGAAGCTTTGCTTGACACACTGGAAGGAGGCTTTCAATGCCAC 1111  
QY 1002 AGCTGTGTGACACATGAGGAAGCTACATGAACTGCACAGACCCCGCGCTC 1056  
DB 1112 CTTGTTCTGCGCCACATCCGGAAGCTGGGCGAGATCCCAAGAGGCGAGGGGCC 1166

RESULT 2

US-09-272-796-12  
; Sequence 12, Application US/09272796  
; Patent No. 6207148  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl G.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
; TITLE OF INVENTION: KINASES  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Incyte Pharmaceuticals, Inc.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/272.796  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/878,989  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0321 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1282 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PROSNOT06  
; CLONE: 827431  
US-09-272-796-12

Query Match 17.1%; Score 418; DB 4; Length 1282;  
Best Local Similarity 65.5%; Pred. No. 2.8e-109;  
Matches 626; Conservative 0; Mismatches 326; Indels 3; Gaps 1;

QY 105 GAAGAAACAGACACCAACATCCGGAAACCTTCATTTTATGGAAGTGTGGGATCAGG 164  
DB 212 GAAGAAACACACGGAGGAGATCAGCAGCGTCTACGAGATCCCGAGAGGCTCGGCTCGG 271  
QY 165 ACCTTTCTCAGAAGTTTCCCTGCTGGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAA 224  
DB 272 TGCCTTCTCCGAGGTGGTGTGGCCAGGAGGCGGCTCCGACACCTCTGGGCCCTCAA 331  
QY 225 GTGCAT---CAAGAAGTCCACCTGCTCCGGACAGCAGCTCGAGAAATGAGATTGCTGT 281  
DB 332 GTGCATCCCAAGAAGGCCCTCCGGGCAAGGAGGCCCTGTGTGGAACGAGATCGCAGT 391  
QY 282 GTTGAAGAAGATCAAGCATGAAACATTTGTGACCCCTGGAGGACATCTATGAGACACCA 341  
DB 392 GCTCGTAGGATCAGTACCCCAACATCGTCGCTCGGAGGATGTCCACGAGAGCCCTTC 451  
QY 342 CCACCTACTACCTGGTGCAGCTGTGTTCTGTTGGGAGCTCTTTGACCGGATCCTTGA 401

Db 452 CCACCTTACCTGGCCATGGAACCTGGTGACGGGTGCGGAGCTGTTTGACCGCATCATGGA 511  
QY 402 GCGGGGTGCTACACAGAGAGATGCCAGTCTGGTGATCCACGAGGCTTTGTCGGCAGT 461  
Db 512 GCGGGGTGCTACACAGAGAGATGCCAGGATCCAGCCATCTGGTGGTCCCTTTGGCGCGT 571  
QY 462 GAAATACCTACATGAGAAATGGCATCTCCACAGAGACTTAAAGCCCGAAACCTGCTTTA 521  
Db 572 CTCCTACCTGCACAGCTGGGATCTGCACCGGACCTCAAGCCGAAACCTCTGTA 631  
QY 522 CTTTACCCCTGAAGAACTCTAAGATCATGATCACTGACTTTGGTCTCTCAAGATGA 581  
Db 632 TGGCCAGCCCTTTGAGGACTCGAAGATCATGTGCTCTGACTTTGGACTCTCCAAAATCCA 691  
QY 582 ACAGATGGCATCATCTGCACCTGCTGCGACCCAGCCTAGTGGCTCCAGAAGTCT 641  
Db 692 GCGTGGAAACATCTAGGACCCGCTGGGACCCCTGGATATGTGGCCCGAGAGCTCTT 751  
QY 642 GGCCCAAAACCTACAGCAAGGCTGTGGATTGCTGGTCCATCGGCGTCACTACCTACAT 701  
Db 752 GGAGCAGAAACCTACGGGAAGGCGTAGATGTGGGCGCTGGGCGTCACTCTCTACAT 811  
QY 702 ATTGCTCTGTGATACCCCGCTTCTATGAAGAAAGGAGTCTAAGCTTTTCGAGAAGAT 761  
Db 812 CTGCTGTGTGGTACCCCGCTTCTACGACGAGAGCGACCTGAGCTCTTCAGCCAGAT 871  
QY 762 CAAGGAGGCTACTATGAGTTGAGTCTCCATCTCGGATGACATTTCTGAGTCAAGCAA 821  
Db 872 CTGAGGGGCGAGTATGAGTTGAGTCTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 931  
QY 822 GGACTTTATTGCACTGCTTGAGAAAGATCCGAACGAGCGGTACACCTGTGAGAAGGC 881  
Db 932 AGACTTTATTGCACTTCTTGAGGAGAGACCTTCAGAAAGAGTTTCACTGCCAACAGGC 991  
QY 882 CTTGAGTCACTGCTGAGTTCAGGAAACACGCGCCCTCCACCGGGGACATCTACCCATCAT 941  
Db 992 CTTGCGGGACCTTTGGATCTTTTGGGACACAGGCTTTGGCAGGACATCTTAGGGTTGT 1051  
QY 942 CAGCCTCCAGATCCAGAAACTTTGCTAAGACGAGTGGAGGCGAGCCCTTCAACGCGAGC 1001  
Db 1052 CAGTGACGATCCGGAAGAACTTTGCTTGGACACACTTGAACGAGGCTTCAATGCCAC 1111  
QY 1002 AGCTGTGTGCACCATCAGGAAGCTACACATGAACCTGACAGCGCGGGGGTCC 1056  
Db 1112 CTTGTTCTTGGCCACATCCGNAAGCTGGGGCAGATCCAGAGGGCGAGGGGGCC 1166

RESULT 3

US-08-905-223-89

; Sequence 89, Application US/08905223

; Patent No. 622029

; GENERAL INFORMATION:

; APPLICANT: Edwards, Jean-Baptiste D.

; APPLICANT: Duclert, Aymeric

; APPLICANT: Lacroix, Bruno

; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS

; NUMBER OF SEQUENCES: 503

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson & Bear

; STREET: 501 West Broadway

; CITY: San Diego

; STATE: California

; COUNTRY: USA

; ZIP: 92101-3505

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: Win95

; SOFTWARE: Word

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/905,223

; FILING DATE:

; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 89:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 362 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: DOUBLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: CDNA  
; ORIGINAL SOURCE:  
; ORGANISM: Homo Sapiens  
; TISSUE TYPE: Brain  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: 87..191  
; IDENTIFICATION METHOD: Von Heijne matrix  
; OTHER INFORMATION: score 5.6  
; OTHER INFORMATION: seq FIFMEVLGSAFS/EV  
US-08-905-223-89

Query Match 13.8%; Score 337.2; DB 4; Length 362;

Best Local Similarity 99.7%; Pred. No. 1.5e-86;

Matches 337; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGAGTGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCCTCAGAGCTTCAACT 60

Db 18 TGGAGTGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCCTCAGAGCTTCAACT 77

QY 61 CTGGAGGCAATGGGTGGAAGAGAGATGACTGCAGTTCCTCGAAGAAACAGACCACC 120

Db 78 CTGGAGGCAATGGGTGGAAGAGAGATGACTGCAGTTCCTCGAAGAAACAGACCACC 137

QY 121 AACATCCGGAAACCTTCATTTTATGGAAGTGTGGATCAGGAGCTTCTCAGAAGTT 180

Db 138 AACATCCGGAAACCTTCATTTTATGGAAGTGTGGATCAGGAGCTTCTCAGAAGTT 197

QY 181 TTCTGTTGGAAGAAAGACTGACTGGGAAGCTTTTCTCTGGAAGTGCATCAAGACTCA 240

Db 198 TTCTGTTGGAAGAAAGACTGACTGGGAAGCTTTTCTCTGGAAGTGCATCAAGACTCA 257

QY 241 CCTGCCTCCGGGACAGCAGCTGGAGAATGAGATTGCTGTGTTGAAAAAGATCAAGCAT 300

Db 258 CCTGCCTCCGGGACAGCAGCTGGAGAATGAGATTGCTGTGTTGAAAAAGATCAAGCAT 317

QY 301 GAAACATTTGACCTGGAGGACATCTATGAGAGCAC 338

Db 318 GAAACATTTGACCTGGAGGACATCTATGAGAGCAC 355

RESULT 4

US-09-734-030-1

; Sequence 1, Application US/09734030

; Patent No. 6461846

; GENERAL INFORMATION:

; APPLICANT: BEASLEY, Ellen M.

; APPLICANT: MERKLOV, Gennady

; APPLICANT: KETCHUM, Karen A.

; APPLICANT: WEI, Ming-Hui

; APPLICANT: DIFRANCESCO, Valentina

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: CLO00612

; CURRENT APPLICATION NUMBER: US/09/734,030

; CURRENT FILING DATE: 2000-12-12

; PRIOR APPLICATION NUMBER: 60/207,281

```

RESULT 5
US-08-715-568A-2
; Sequence 2, Application US/08715568A
; Patent No. 5856463
; GENERAL INFORMATION:
; APPLICANT: Prydz, Hans Peter Blankenborg
; APPLICANT: Brede, Gaute
; TITLE OF INVENTION: PSKH-1 Ribozymes and Uses in Disease
; TITLE OF INVENTION: Treatment
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lerner, David, Littenberg, Krumholz & Mentlik
; STREET: 600 South Avenue West
; CITY: Westfield
; STATE: NJ
; COUNTRY: USA
; ZIP: 07090-1497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/715,568A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Foley, Shawn P.
; REFERENCE/DOCKET NUMBER: FORSK 3.0-002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-654-5000
; TELEFAX: 908-654-7866
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3471 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-715-568A-2

Query Match      8.1%; Score 197.2; DB 2; Length 34
Best Local Similarity 54.3%; Pred. No. 4.8e-46;
Matches 425; Conservative 0; Mismatches 348; Indels

Qy 143 TTATGGAAGTGTGGGATCAGGAGCTTTCTCTACAAGTTTTCCTGGTGAAGCAAA
Db 393 TCAAGGAATAATTGCCCGAGGAGCAGCTTCAGCCGAGTGGTACGTGTAGAGGAC
Qy 203 CTGGGAAGCTCTTTCCTCTGAAGTGCATCACAAGATCAGCTCCCTTCCTCGGGAC
Db 453 CCCGCAACCCGTATGCCATCAAGATGATGTAGACCAAGTACC CGGGAGGCGCGGG
Qy 263 TGGAGAANTGAGATTCTGTTGTGAAAAAGATCAAGCATGAAACATTTGTGACCC
Db 513 GTGAGTCGGAGCTGCGTGTGCTGCGTCCGTGCGGTCTATGCCAACATCATCCAGC
Qy 323 ACATCTATGAGAGCACCACCCACTACTACCTGGTCATGACAGCTTGTTCCTGGTG
Db 573 AGGTGTTCCGAGACACAGGAGCGGGGTGTACATGGTATGGAGCTGGCCACTGTGTG
Qy 383 TCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAGNAGGATGCCAGTCTGG
Db 633 TCTTTGACCGGATCAATTGCCAAGGGCTCCTTCACGAGCGGTGACGCCACGCGGG
Qy 443 AGCAGGCTCTTGTGGCAGTGGAAATACCTACATGGAATGGCATCGTCCACAGAG
Db 693 AGATGGTCTGGATGSGCTGCCGGTATCTGCATGSCATGGGCATCACACACCGAG
Qy 503 AGCCCGAAAACTGCTTTTACCTTACCCCTGAAGAGAACTCTTAAGATCATGATCA
Db 753 AACCTGAGAAATCTGCTTACTACCTACCATCCGGGCACTGACTCCAAAGATCATCATCA

```

QY 563 TTGCTGTGTC-----CAAGATGAACAGAGATGGCATCATGTCCTGCTGGGA 613  
Db 813 TCGGCTTGGCCAGTGTGCGCAAGAGGGTGATGACTGTGATGAAGACCACTGTGGCA 872  
QY 614 CCCAGGCTACGTGGCTCCAGAAAGTGTGCGCCAGAAACCCCTACAGCAAGGCTGTGGATT 673  
Db 873 CGCCTGAGTACATTGCCCGAGAGTCTGTGTCGGAAGCCATACACCACTCAGTGGACA 932  
QY 674 GCTGGTCCATCGGCGTACATCACCACATATGCTCTGTGGATACCCCGGTTCTATGAAG 733  
Db 933 TGTGGCGCTGGGCGTCACTTCCCTACATCTACTAGTGGCACCATGCCGTTTGAGGATG 992  
QY 734 AAGCGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTACTATGAGTTTGAGTCTCCAT 793  
Db 993 ACACCGGTACCCGGCTGTACCGGCAGATCCTCAGGGGCAAGTACAGTTACTCTGGGGAGC 1052  
QY 794 TCTGGGATGACATTTCTGAGTCAGCAAGGACTTTATTTGCCACATGCTGTGAGAAGGATC 853  
Db 1053 CTTGGCTAGTGTGTCACACCTGGCCAGGACTTCATTGACCGCTGCTGACAGTGGACC 1112  
QY 854 CGAACAGCGGTACACCTGTGAGAAGCGCTTGAGTCATCCCTGGATGACGGAAACACGG 913  
Db 1113 CTGAGCGCGTATGACTGCACTCAGGCGCTGAGGACCCGCTGGGTGAGCATGGCTG 1172  
QY 914 CC 915  
Db 1173 CC 1174

RESULT 6  
US-09-800-960-1  
; Sequence 1, Application US/09800960  
; Patent No. 6387677  
; GENERAL INFORMATION:  
; APPLICANT: YE, Jane et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001158  
; CURRENT APPLICATION NUMBER: US/09/800,960  
; CURRENT FILING DATE: 2001-03-08  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2061  
; TYPE: DNA  
; ORGANISM: Human  
US-09-800-960-1

Query Match 6.9%; Score 168.6; DB 4; Length 2061;  
Best Local Similarity 54.2%; Pred. No. 5.2e-38;  
Matches 367; Conservative 0; Mismatches 304; Indels 6; Gaps 1;

QY 229 ATCAAGAAGTCAGCTGCTTCCGGGACAGAGCTGGGAATGAGATGCTGTGTGAAA 288  
Db 251 ACCAAGAAATTGTCGCCGGGATCACAGAACTAGACGTGAGCTCGGATATGCGA 310  
QY 289 AAGATCAAGCATGAACAATGTGACCCCTGGAGGACATCTATCAGAGCACCAACCCACTAC 348  
Db 311 CTTCTGAACATCCAAACATCGTGCCTCCATGACAGATTTCTGAAGAAGGGTTTTCAC 370  
QY 349 TACTGTGTATCAGCTGTTTCTGTGGGAGCTCTTTGACGGGATCTCTGGAGCGGGT 408  
Db 371 TACTCTGTTGACCTGTGTACCGGGGGAGCTGTTTGAAGACATTTGCGCGAGAG 430  
QY 409 GTCTACAGAGAAGGATGCCAGTCTGGTGATCCAGCAGGCTCTGTGCGCAGTGAATAAC 468  
Db 431 TACTACAGTCAAGCAGATGCCACCACCTGTATACATCAGATTCTGGAGAGTGTAAACCAC 490  
QY 469 CTACATGAGAAATGGCTGCTGACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACC 528  
Db 491 ATCCACCAGCATGACATCGTCCACAGGGACCTGAAGCCTTGAGAACCTGTGCTGGCGAGT 550

QY 529 CCTGAGAGAAGCTCTAAGATCATCATGACTGCTTGGTCTGTCCTCAAGATGGACAGAAT 588  
Db 551 AAATGCAAGGTGCGCGCTCAAGCTGGCTGATTTTGGCTTAGCCATCGAAGTACAGGA 610  
QY 589 GGCATCATGTGCCA-----CTGGCCTGTGGGACCCAGGCTACGTGGCTCCAGAAAGTCTG 642  
Db 611 GAGCAGCAGGCTGTGGTTTGGTTTGGCTGGCACCCAGGTTACTTGTCCCTGAGGCTTG 670  
QY 643 GCCCAGAAACCTTACAGCAGGCTGTGATTTGGTTCATCGGCGTCACTACACATACATA 702  
Db 671 AGGAAAGATCCTATGAAACACCTGTGATATCTGGGCTCGGGGTTCATCTGTATATC 730  
QY 703 TTGCTCTGTGATACCCCGCTTCTATGAAGAACGAGTCTAGCTTTTCGAGAAATC 762  
Db 731 CTCTGTGGGCTATCTCTCCCTTCTGGGATGAGGATCAGCACAAAGCTGTATCAGCAGATC 790  
QY 763 AAGGAGGCTACTATGAGTTTGGCTCTCCATTCCTGGGATGACATTTCTGAGTCAGCCAAG 822  
Db 791 AAGCTGAGGCTATGATTTCCCATCACCAGATGGGACACGGTAACTCCTGAAGCCAAG 850  
QY 823 GACTTTATTTCCCACTTGTGAGAAGATCCGAACGAGCGGTACACCTGTGAGAAGGCC 882  
Db 851 AACTTGATCAACCAAGATGCTGACCATAAACCCAGCAAGGCGCATCAGGCTGACCGGCT 910  
QY 883 TTGAGTCATCCCTGGAT 899  
Db 911 CTCAAGCACCCGTGGGT 927

RESULT 7  
US-09-159-385-3  
; Sequence 3, Application US/09159385  
; Patent No. 5958748  
; GENERAL INFORMATION:  
; APPLICANT: KAWAI, TARO  
; APPLICANT: AKIRA, SHIZUO  
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE  
; FILE REFERENCE: PH-569  
; CURRENT APPLICATION NUMBER: US/09/159,385  
; CURRENT FILING DATE: 1998-09-23  
; EARLIER APPLICATION NUMBER: JP97/261589  
; EARLIER FILING DATE: 1997-09-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2132  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (94)..(1455)  
US-09-159-385-3

Query Match 6.7%; Score 163.4; DB 2; Length 2132;  
Best Local Similarity 55.3%; Pred. No. 1.6e-36;  
Matches 361; Conservative 0; Mismatches 286; Indels 6; Gaps 2;

QY 263 TGGAGATGAGATTGCTGTGTGTGAAAAGATCAAGCATGAAACATTTGTGACCTGAGG 322  
Db 275 TCGAGCGGGAGGTGAACATCTCTGGGGAGATCCGGCACCCCAACATCATCACCCTGCACG 334  
QY 323 ACATCTATGAGAGCACCAACCCACTACTACCTGGTCTGTCAGCTGTGTTCTGGTGGGAGC 382  
Db 335 ACATCTCGAGAACAAAGACGACGCTGCTTCCTCTGAGGAGTGTCTGCGGGGAGC 394  
QY 383 TCTTTGACCGGATCCTGGAGCGGGGTGTCTATACAGAGAAGGATGCCAGTCTGTGTATCC 442  
Db 395 TCTTTGACTTCTTGGCGGAGAAAGAGTGGCTGACGAGGAGGAGGCCACCCAGTTCCTCA 454  
QY 443 AGCAGCTTTGTGGCAGTGAATACCTACATGAGATGGCATCGTCCACAGAGACTTAA 502  
Db 455 AGCAGATCCTGGAGCGGCTTCACTACCTGCACTCTAAGCGCATCGCACACTTTGACCTGA 514





NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP  
TELEPHONE: (919)541-8615  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1349 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..1226  
OTHER INFORMATION: /note= "cDNA sequence for maize  
OTHER INFORMATION: pollen-specific calcium dependent protein kinase gene as  
disclosed in Figure 30."  
US-07-951-715A-20

Query Match 5.5%; Score 135.6; DB 1; Length 1349;  
Best Local Similarity 51.5%; Pred. No. 1.1e-28;  
Matches 312; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 304 AACATTGTGACCTGGAGGACATCTATGAGAGACACCCACTACTACTCGTGTCTATGAGAGAG 363  
DB 33 AACGTTGGTGGGCTCCGCGGCGGTACGAGGACAGACGCGTGCACCTCGTCTATGAG 92  
QY 364 CTTTGTCTGTGGGAGCTCTTTGACCGGATCTTGGAGCGGGTGTCTACACAGAGAA 423  
DB 93 CTGTGCGGCGGGGAGCTCTTGACCGCATCATCGCGCGGCCGACACACGAGCGC 152  
QY 424 GATGCCAGTCTGTGATCCAGAGGTCTTGTGCGAGTGAATACCTACATGAGAATGGC 483  
DB 153 GCGCGCGGGAGCTCTCGCGGCCATCGTGCAGATCTGTGCACACTGCCACTCCATGGG 212  
QY 484 ATCGTCCACAGAGACTTAAAGCCGAAACCTGTTTACCTTACCCCTGAAGAGAACTCT 543  
DB 213 GTGATGACCGGGACATCAAGCCGAGAACTTCTGTGCTCAGCAGGAGGAGCGCG 272  
QY 544 AAGATCATGATCACTGACTTTTGGTGTGTCCAGATGGAACAGATGGCATCATGTCCACT 603  
DB 273 CGGCTCAAGGCGACCGACTTGGCGCTCTCCGCTCTTCTTCAAGGAGGCGAGCTGCTCAGG 332  
QY 604 GCCTGTGGGACCCAGGCTAGTGGCTCCAGAAGTGTGGCCGACCAACCTACAGCAAG 663  
DB 333 GACATCGTGGGAGCGCTACTACATCGCGCGGAGGTGCTCAAGAGGAAGTACGGCCG 392  
QY 664 GCTGTGGATTGCTGTCCATCGGCTCATCACCTACATPATTTGCTGTGGATACCCCGC 723  
DB 393 GAGGCGGACATCTGGAGCGTGGCGTCTATGCTTACATCTTCTTCCGCGGCGTGCCTCC 452  
QY 724 TTCTATGAAGAACGGAGTCTAAGCTTTTCGAGAGATCAAGGAGGCTACTATGAGTTT 783  
DB 453 TTCTGGGAGAGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 512  
QY 784 GAGTCTCCATTCTGGGATTGAGTTCCTGAGTCCAGCAGGAGGAGTTCATTTGGCACTTGCTT 843  
DB 513 TCCAGGAGCCATGCCACACATCTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 572  
QY 844 GAGAGGATCCGAGCAGCGGTACACCTGTGAGAGGCGCTTTGAGTCACTCCCTGGATTGAC 903  
DB 573 AACATCAACCCCAAGGAGGCGGTACGCGGTTCAGGCTTCCAGGCTCTCAATCACCCATGATCAAA 632  
QY 904 GGAAC 909  
DB 633 GAAGAC 638

RESULT 12  
US-08-459-448A-20

Sequence 20, Application US/08459448A  
Patent No. 5859336  
GENERAL INFORMATION:  
APPLICANT: Koziel, Michael G.  
APPLICANT: Desai, Nalini M.  
APPLICANT: Lewis, Kelly S.  
APPLICANT: Kramer, Vance C.  
APPLICANT: Warren, Gregory W.  
APPLICANT: Evola, Stephen V.  
APPLICANT: Crossland, Lyle D.  
APPLICANT: Wright, Martha S.  
APPLICANT: Merlin, Ellis J.  
APPLICANT: Launis, Karen L.  
APPLICANT: Rothstein, Steven J.  
APPLICANT: Bowman, Cindy G.  
APPLICANT: Dawson, John L.  
APPLICANT: Dunder, Erik M.  
APPLICANT: Pace, Gary M.  
APPLICANT: Suttie, Janet L.  
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED  
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 5859336artis Corporation  
STREET: Patent & Trademark Dept., 520 White Plains  
STREET: Rd., POB 2005  
CITY: Tarrytown  
STATE: New York  
COUNTRY: USA  
ZIP: 10591-9005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,448A  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/951,715  
FILING DATE: 25-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/772,027  
FILING DATE: 04-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Pace, Gary M.  
REGISTRATION NUMBER: 40403  
REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIVA  
TELEPHONE: (919)541-8582  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1349 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..1226  
OTHER INFORMATION: /note= "cDNA sequence for maize  
OTHER INFORMATION: pollen-specific calcium dependent protein kinase gene as  
disclosed in Figure 30."  
US-08-459-448A-20

Query Match 5.5%; Score 135.6; DB 2; Length 1349;  
Best Local Similarity 51.5%; Pred. No. 1.1e-28;  
Matches 312; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 304 AACATTGTGACCTGGAGGACATCTATGAGAGCACCACCCACTACTACTCGTGTCTATGAG 363

Db 33 AACGTGGTGGCCCTCCGCGCGGTACGAGACAAGCAGACGCTGCGACCTCGTCATGGAG 92  
Qy 364 CTTGTTCTGTGGGAGCTCTTTGACCGATCTTGAGCGGGGTGCTACACAGAGAAG 423  
Db 93 CTGTGCGCGGGGGAGCTCTTGACCGATCATCGCCGCGCCAGTACACGGAGCGC 152  
Qy 424 GATGCACTGTGGTACCAAGCAGGTCTTGTGCGCAGTGAATACCTACATGAGATGCC 483  
Db 153 GCGCGCGGAGCTGCTGCGGCGCATGTCGATGTCGACACCTGCCACTCCATGGGG 212  
Qy 484 ATCGTCCACAGAGACTTAAAGCCGAAACCTGCTTACCTTACCCCTGAAGAGAATCT 543  
Db 213 GTGATGACCGGACATCAAGCCGAGAACTTCTGCTGCTCAGCAAGGAGGAGCGC 272  
Qy 544 AAGATCATGATCATGACTTTGGTCTGCTCAAGATGGAACAGATGGCATGTCCTACT 603  
Db 273 CCGCTCAAGGCCACCGACTTGGGCTCTCCGCTCTTCTTCAAGGAGGCGGAGCTGCTCAGG 332  
Qy 604 GCCTGTGGGACCCAGGCTAGTGGCTCCAGAAAGTGTGCGCCGCAAAACCTACAGCAAG 663  
Db 333 GACATGTCGCGACGCGCTACTACATCGCGCCGAGGTGCTCAAGAGGAAGTACGCCCG 392  
Qy 664 GCTGTGATGCTGGTCCATCGGCGTCATCACCTACATATGCTGTGATACCCCGCG 723  
Db 393 GAGCGCGACATCTGGAGCGTGGCGTCATGCTCTACATCTTCTCGCGCGCGTCTGCC 452  
Qy 724 TTCTATGAAGAAGGAGTCTAAGCTTTTCGAGAAGTCAAGAGGCGCTACTAGATT 783  
Db 453 TTCTGGCAGAGAAGCAGACGCGATCTTCACCGCCCATCTCGGAGGCGGAGCTGACCTC 512  
Qy 784 GAGTCTCCATTTCTGGATGACATTTCTGAGTCAGCAAGGACTTTATTTGCCACTTGC 843  
Db 513 TCCAGCGAGCATGGCCACACATCTCGCGGGAGCAAGGATCTGTCARAAGATGCTC 572  
Qy 844 GAGAAGATCGAACGAGCGGTACACTGTGAGAAGCGCTTGAGTCATCCCTGGATTGAC 903  
Db 573 AACATCAACCCCAAGGAGCGCTCACGCGCTTCCAGGTCTCTCAATCACCCATGATCAA 632  
Qy 904 GGAAC 909  
Db 633 GAAGAC 638

RESULT 13

US-08-459-595A-20  
; Sequence 20, Application US/08459595A

; Patent No. 6018104

GENERAL INFORMATION:

; APPLICANT: Koziel, Michael G.  
; APPLICANT: Desai, Nalini M.  
; APPLICANT: Lewis, Kelly S.  
; APPLICANT: Kramer, Vance C.  
; APPLICANT: Warren, Gregory W.  
; APPLICANT: Evola, Stephen V.  
; APPLICANT: Crossland, Lyle D.  
; APPLICANT: Wright, Martha S.  
; APPLICANT: Merlin, Ellis J.  
; APPLICANT: Launis, Karen L.  
; APPLICANT: Rothstein, Steven J.  
; APPLICANT: Bowman, Cindy G.  
; APPLICANT: Dawson, John L.  
; APPLICANT: Dunder, Erik M.  
; APPLICANT: Pace, Gary M.  
; APPLICANT: Suttie, Janet L.  
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED  
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE  
; NUMBER OF SEQUENCES: 94  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 6018104artis Corporation  
; STREET: Patent & Trademark Dept., 520 White Plains  
; STREET: Rd., POB 2005  
; CITY: Tarrytown

; STATE: New York  
; COUNTRY: USA  
; ZIP: 10591-9005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/459,595A  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/951,715  
; FILING DATE: 25-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/772,027  
; FILING DATE: 04-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pace, Gary M.  
; REGISTRATION NUMBER: 40403  
; REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (919)541-8582  
; TELEFAX: (919)541-8689  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1349 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 3..1226  
; OTHER INFORMATION:  
; OTHER INFORMATION: pollen-specific calcium dependent protein kinase gene as  
; OTHER INFORMATION: disclosed in Figure 30."  
; US-08-459-595A-20

Query Match 5.5%; Score 135.6; DB 3; Length 1349;

Best Local Similarity 51.5%; Pred. No. 1.1e-28;

Matches 312; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 304 AACATGTGACCCCTGGAGGACATCTATGAGAGCACCAACCACTACTACTGCTCATGCG 363  
Db 33 AACGTGGTGGGCTCCGCGCGGTACGAGACAAGCAGACGCTGCTCATGGAG 92  
QY 364 CTTGTTCTGTGGGAGCTCTTTGACCGGATCTTGGAGCGGGGTGCTACACAGAGAAG 423  
Db 93 CTGTGCGGGGGGGAGCTCTTTCAGCCGCATCATCGCCGCGGCGCAGTACACGGAGCGC 152  
QY 424 GATGCACTGTGGTATCCAGCAGGTCTTGTGCGCAGTGAATACCTACATGAGAAATGCC 483  
Db 153 GCGCGCGGAGCTGCTGCGCGCCATGTCGAGATCGTGCACACTGCCACTCCATGGGG 212  
QY 484 ATCGTCCACAGAGACTTAAAGCCGAAACCTGCTTACCTTACCCCTGAAGAGAATCT 543  
Db 213 GTGATGACCGGACATCAAGCCGAGAACTTCTGCTGCTCAGCAAGGAGGAGCGC 272  
QY 544 AAGATCATGATCATGACTTTGGTCTGCTCAAGATGGAACAGATGGCATGTCCTACT 603  
Db 273 CCGCTCAAGGCCACCGACTTGGGCTCTCCGCTCTTCTTCAAGGAGGCGGAGCTGCTCAGG 332  
QY 604 GCCTGTGGGACCCAGGCTAGTGGCTCCAGAAAGTGTGCGCCGCAAAACCTACAGCAAG 663  
Db 333 GACATGTCGCGACGCGCTACTACATCGCGCCGAGGTGCTCAAGAGGAAGTACGCCCG 392  
QY 664 GCTGTGATGCTGGTCCATCGGCGTCATCACCTACATATGCTGTGATACCCCGCG 723  
Db 393 GAGCGCGACATCTGGAGCGTGGCGTCATGCTCTACATCTTCTCGCGCGCGTCTGCC 452



Warren, Gregory W.  
Evola, Stephen V.  
Crossland, Lyle D.  
Wright, Martha S.  
Merlin, Ellis J.  
Launis, Karen L.  
TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED  
NUCLEIC ACID CODING SEQUENCE  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/459,444A  
FILING DATE: 02-Jun-1995  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/951,715  
FILING DATE: 25-SEP-1992  
APPLICATION NUMBER: US 07/772,027  
FILING DATE: 04-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: S-18805/P1/CGC1577/CIP/DIV6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8587  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: /note= "cDNA sequence for maize  
pollen-specific calcium dependent protein kinase gene as  
disclosed in Figure 30."

SEQUENCE CHARACTERISTICS:  
LENGTH: 1349 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..1226  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-08-459-444-20

Query Match 5.5%; Score 135.6; DB 3; Length 1349;  
Best Local Similarity 51.5%; Pred. No. 1.1e-28;  
Matches 312; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

Qy	304	AACATTGTGACCTGGAGGACATCTATGAGAGCACCACCCTACTACTCTGGTGCATGCAG	363
Db	33	AACGTGTGGTCCGCGCGGTACGAGCAAGCAGCGTGCACCTCGTCATGGAG	92
Qy	364	CTTGTGTTCTGGTGGGAGCTTTTGACCGGATCCTGGAGCGGTGTCTACACAGAGAAG	423
Db	93	CTGTGCGCGGGGGGAGCTCTTCGACCGCATCATCGCCGGGCCAGTACACGGAGCGC	152
Qy	424	GATGCCAGTCTGGTGATCCAGCAGGCTTTGTCGGCAGTGAATACCTACATGAGATGGC	483
Db	153	GGCGCCGCGAGCTGTGCGCGCATCGTGCAGATCGTGCACACCTGCCACTCCATGGGG	212
Qy	484	ATCGTCCACAGACATTAAAGCCGCAACCTGCTTTACCCCTGAAGAGAACTCT	543
Db	213	GTGATGACCGCGGACATCAAGCCCGAGACTTCTCTGCTCTCAGCAAGGACGAGCGG	272

Qy	544	AAGATCATGATCACTGACTTTTGGTCTCTTCAAGATGGAACAGATGGCATCATGTCCACT	603
Db	273	CCGCTCAAGGCCACCGACTTTCGGCTCTCTCGTCTCTTCAAGGAGGCGGAGCTCTCAGG	332
Qy	604	GCCTGTGGGACCCAGGCTAGTGGCTCCAGAAAGTGTGGCCCAAGAAACCCCTACAGCAAG	663
Db	333	GACATCGTTCGGCAGCGCTACTACATCGCGCCCGAGGTGCTCAAGAGGAAGTACGGCCCG	392
Qy	664	GCTGTGGATTGCTGCTCCATCGGCTCATCACCTACATATTGCTCTCTGGATACCCCCCG	723
Db	393	GAGCCGACATCTGGAGGTCGGCGTCTATGCTCTACATCTCTCTCGCGGGTGCCTCCC	452
Qy	724	TTCTATGAAGAAACGGAGTCTTAAGCTTTTCGAGAAGATCAAGAGGCGCTACTATGAGTTT	783
Db	453	TTCTGGGCAGAGAACGAGACGGCATCTTCACCGCCATCTCTGCGAGGCGAGCTTGACCTC	512
Qy	784	GAGTCTCCATTCTGGGATGACATTTCTGAGTCAAGCAAGGACTTTATTTGCCACTTCTT	843
Db	513	TCCAGCGAGCCATGGCCACACATCTCGCCGGGAGCCCAAGGATCTCGTCAAGAAGATGCTC	572
Qy	844	GAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGGCTTTGAGTCAATCCCTGGATTGAC	903
Db	573	AACATCAACCCCAAGGAGCGGCTCAGCGCTTCCAGGTCTCTCAATCAATCAATGATCAAA	632
Qy	904	GGAAGAC 909	
Db	633	GAAGAC 638	

Search completed: March 14, 2003, 17:23:14  
Job time : 102 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:07:56 : Search time 16 Seconds  
(without alignments)  
875.332 Million cell updates/sec

Title: US-09-960-643-2  
Perfect score: 2513  
Sequence: 1 MGRKEEDDCSSWKQFTNIR.....VKAGSSHCRAQTGVCLIM 476

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1152.5	45.9	370	2	US-08-878-989-19
2	1152.5	45.9	370	4	US-09-272-796-19
3	1152.5	45.9	370	4	US-09-457-0408-31
4	990	39.4	343	2	US-08-878-989-5
5	990	39.4	343	4	US-09-272-796-5
6	747.5	29.7	501	4	US-09-734-030-2
7	650.5	25.9	424	2	US-08-715-568A-1
8	625	24.9	264	2	US-07-857-224B-24
9	592	23.6	556	4	US-09-800-960-4
10	587.5	23.4	264	2	US-07-857-224B-18
11	578.5	23.0	543	4	US-09-529-093A-2
12	571.5	22.7	565	4	US-09-800-960-2
13	564	22.4	264	2	US-07-857-224B-19
14	563.5	22.4	295	1	US-07-951-715A-23
15	563.5	22.4	295	2	US-08-459-448A-23
16	563.5	22.4	295	3	US-08-459-595A-23
17	563.5	22.4	295	3	US-08-459-504B-23
18	563.5	22.4	295	3	US-08-459-444-23
19	563.5	22.4	295	4	US-09-547-422-23
20	558	22.2	776	1	US-08-198-446B-17
21	558	22.2	776	2	US-08-870-693-17
22	558	22.2	821	1	US-08-198-446B-6
23	558	22.2	821	2	US-08-870-693-6
24	542.5	21.6	463	1	US-07-951-715A-25
25	542.5	21.6	463	2	US-08-459-448A-25
26	542.5	21.6	463	3	US-08-459-595A-25
27	542.5	21.6	463	3	US-08-459-504B-25

28	542.5	21.6	463	3	US-08-459-444-25
29	542.5	21.6	463	4	US-09-547-422-25
30	538.5	21.4	639	4	US-09-347-801-17
31	537.5	21.4	456	1	US-08-464-164-2
32	537.5	21.4	456	1	US-08-338-057-2
33	537.5	21.4	456	2	US-08-668-416-2
34	531	21.1	520	4	US-09-257-825B-20
35	530	21.1	454	2	US-09-159-385-1
36	530	21.1	454	4	US-09-186-277-1
37	527.5	21.0	382	4	US-09-142-551A-2
38	525.5	20.9	1423	4	US-08-810-712-10
39	522	20.8	448	2	US-09-159-385-2
40	522	20.8	448	4	US-09-186-277-2
41	520.5	20.7	331	4	US-08-810-712-24
42	510	20.3	517	4	US-09-257-825B-21
43	498	19.8	260	2	US-07-857-224B-23
44	496.5	19.8	625	4	US-09-347-801-18
45	495.5	19.7	623	4	US-09-347-801-4

ALIGNMENTS

RESULT 1  
US-08-878-989-19  
; Sequence 19, Application US/08878989  
; Patent No. 5885803  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl G.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/878,989  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0321 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 370 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank

```
; CLONE: 790790
US-08-878-989-19

Query Match 45.9%; Score 1152.5; DB 2; Length 370;
Best Local Similarity 55.8%; Pred. No. 1.5e-93;
Matches 239; Conservative 46; Mismatches 62; Indels 81; Gaps 9;

QY 12 WKQOTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPAFRDSLENEIA 70
Db 10 W-KQAEIRDIDYDFRDLGTGAFSEVILAEDKRTQKLVAKIAKEALEGKSGMENEIA 68

QY 71 VLKIKHENIVTLEDIYESTHYIYLMQVSGGELFDRILRGVYTEKDAISLVIOQVLSA 130
Db 69 VLHKIKHPNIVALDDIYESGGHLYLMQVSGGELFDRIVEKGFYTERDASRLIFQVLSA 128

QY 131 VKYLHENGIVHRDLKPNLLYLTPEENSKIMITDGLSKMEQNG-IMSTACGTPGVVAPE 189
Db 129 VKYLDHGVHRDLKPNLLYSLDEDSKIMISDFGLSKMEDPGSVLSTACGTPGVVAPE 188

QY 190 VLAQKPSKAVDCWSIGVITYILLGYPPEYETESKLFKEIKEGYEFESPFWDIDISES 249
Db 189 VLAQKPSKAVDCWSIGVITYILLGYPPEYETESKLFKEIKEGYEFESPFWDIDISES 248

QY 250 AKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVLSQIQKNFAKSKWQAQFN 309
Db 249 AKDFIRHLMKDPKERTCEQALQHPWIAGDTALDKNIHQSVSEQIKKNFAKSKWQAQFN 308

QY 310 AAUVVHMRKLMNLHSPGVPRPEVENRPPETQASSETSRPSSPEITITEAPVLDHVSALPA 369
Db 309 ATAVVRHMRKLQLG-----TS----- 324

QY 370 LTQLPCOHRRTAPGGRSLNCLVNGSLHSSSLVPMHQGSLAAGPCGCCSCLNIGSKG 429
Db 325 -----QEGQGQASHGELLTPVAGGP-----AAG-C-CRDC----- 354

QY 430 KSSYCSEP 437
Db 355 ----CVEP 358

RESULT 2
US-09-272-796-19
; Sequence 19, Application US/09272796
; Patent No. 6207148
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,796
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/878,989
```

```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 370 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 790790
US-09-272-796-19

Query Match 45.9%; Score 1152.5; DB 4; Length 370;
Best Local Similarity 55.8%; Pred. No. 1.5e-93;
Matches 239; Conservative 46; Mismatches 62; Indels 81; Gaps 9;

QY 12 WKQOTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPAFRDSLENEIA 70
Db 10 W-KQAEIRDIDYDFRDLGTGAFSEVILAEDKRTQKLVAKIAKEALEGKSGMENEIA 68

QY 71 VLKIKHENIVTLEDIYESTHYIYLMQVSGGELFDRILRGVYTEKDAISLVIOQVLSA 130
Db 69 VLHKIKHPNIVALDDIYESGGHLYLMQVSGGELFDRIVEKGFYTERDASRLIFQVLSA 128

QY 131 VKYLHENGIVHRDLKPNLLYLTPEENSKIMITDGLSKMEQNG-IMSTACGTPGVVAPE 189
Db 129 VKYLDHGVHRDLKPNLLYSLDEDSKIMISDFGLSKMEDPGSVLSTACGTPGVVAPE 188

QY 190 VLAQKPSKAVDCWSIGVITYILLGYPPEYETESKLFKEIKEGYEFESPFWDIDISES 249
Db 189 VLAQKPSKAVDCWSIGVITYILLGYPPEYETESKLFKEIKEGYEFESPFWDIDISES 248

QY 250 AKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVLSQIQKNFAKSKWQAQFN 309
Db 249 AKDFIRHLMKDPKERTCEQALQHPWIAGDTALDKNIHQSVSEQIKKNFAKSKWQAQFN 308

QY 310 AAUVVHMRKLMNLHSPGVPRPEVENRPPETQASSETSRPSSPEITITEAPVLDHVSALPA 369
Db 309 ATAVVRHMRKLQLG-----TS----- 324

QY 370 LTQLPCOHRRTAPGGRSLNCLVNGSLHSSSLVPMHQGSLAAGPCGCCSCLNIGSKG 429
Db 325 -----QEGQGQASHGELLTPVAGGP-----AAG-C-CRDC----- 354

QY 430 KSSYCSEP 437
Db 355 ----CVEP 358

RESULT 3
US-09-457-040B-31
; Sequence 31, Application US/09457040B
; Patent No. 6387641
; GENERAL INFORMATION:
; APPLICANT: Vertex Pharmaceuticals Incorporated
; APPLICANT: Bellon, Steve
; TITLE OF INVENTION: Crystallized P38 Complexes
; FILE REFERENCE: VPI/98-14
; CURRENT APPLICATION NUMBER: US/09/457,040B
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent version 3.0
; SEQ ID NO 31
; LENGTH: 370
; TYPE: PRT
```

```
; ORGANISM: Human
; US-09-457-040B-31

Query Match      45.9%; Score 1152.5; DB 4; Length 370;
Best Local Similarity 55.8%; Pred. No. 1.5e-93;
Matches 239; Conservative 46; Mismatches 62; Indels 81; Gaps 9;

QY 12 WKKOTNIRKTFIFMEVLGSGAFSEVFLVKORLTGKLFALKCI-KKSPAFRDSLSLENEIA 70
Db 10 W-KOADIIRIDYDFRDLVLTGAFSEVFLAEADRKTQKLVAKICIAKEALGEGKSGMENETA 68
QY 71 VLAKIKHENIVTLEDIYESTHYLYVMQLVSGGELFDRLERGVYTERKDSALVIOQVLSA 130
Db 69 VLHKIKHPNIVALDDIYESGHLVLYLMQLVSGGELEDRIVEKGYTERDASLIIFQVLD 128
QY 131 VKYLHENGIVHRDLKPNENLLYLPENSKIMITDFGLSKMEQNG-IMSTACGTPGVVAPE 189
Db 129 VKYLDHGIIVHRDLKPNENLLYSLDEDSKIMISDFGLSKMEDPGSVLSTACGTPGVVAPE 188
QY 190 VLAQPKYSKAVDCWSIGVITYILLCGYPPFYETESKLFKEIKEGYEFESPFWDIDISES 249
Db 189 VLAQPKYSKAVDCWSIGVITYILLCGYPPFYDENDAKLFEILKAEYEFDSPIWDDISDS 248
QY 250 AKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQTKNFASKKWRQAFN 309
Db 249 AKDFIRHLMKEKDPKRFCTEQALQHPWIAAGDTALDKNIHQSYSEQLKKNFASKKWRQAFN 308
QY 310 AAIVVHMRKLNHLNLSHPGVRPEVENRPPETQASERTSRPSSPEITITTEAPVLDHVALPA 369
Db 309 ATAVVHMRKLQLG-----TS----- 324
QY 370 LTLPCHGRPRTPAGGRSLNCLVNGSLHTSSSLVPMHQSAAAGCGCCSCLNIGSKG 429
Db 325 -----QEGGQGTASHGELLTPVAGGP-----AAG-C-CRDC----- 354
QY 430 KSSYCSEP 437
Db 355 ----CVEP 358

RESULT 4
US-08-878-989-5
; Sequence 5, Application US/08878989
; Patent No. 5885803
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,989
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

; ORGANISM: Human
; US-09-272-796-5
; Sequence 5, Application US/09272796
; Patent No. 6207148
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

; ORGANISM: Human
; US-08-878-989-5
; Query Match      39.4%; Score 990; DB 2; Length 343;
; Best Local Similarity 59.8%; Pred. No. 2.9e-79;
; Matches 193; Conservative 49; Mismatches 73; Indels 8; Gaps 2;

QY 13 KKOTNIRKTFIFMEVLGSGAFSEVFLVKORLTGKLFALKCI-KKSPAFRDSLSLENEIAV 71
Db 5 KKHEDISSVTEIRERLGSFSEVFLAQRGSAHLVALKICPKKALRGREALVENEIAV 64
QY 72 LKKIKHENIVTLEDIYESTHYLYVMQLVSGGELFDRLERGVYTERKDSALVIOQVLSAV 131
Db 65 LRRISHPNIVALDDIYESGHLVLYLMQLVSGGELEDRIVEKGYTERDASLIIFQVLD 124
QY 132 KYLHENGIVHRDLKPNENLLYLPENSKIMITDFGLSKMEQNGIMSTACGTPGVVAPEVL 191
Db 125 SYLSLGIIVHRDLKPNENLLYATPFEDSKIMVSDFLSKIQAGNNMLGTACGTPGVVAPELL 184
QY 192 AQPKYSKAVDCWSIGVITYILLCGYPPFYETESKLFKEIKEGYEFESPFWDIDISESAK 251
Db 185 EQPKYSKAVDCWSIGVITYILLCGYPPFYETESKLFKEIKEGYEFESPFWDIDISESGK 244
QY 252 DFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQTKNFASKKWRQAFNAA 311
Db 245 DFIRHLLERDLQKRFCTEQALQHPWIAAGDTALDKNIHQSYSEQLKKNFASKKWRQAFNAT 304
QY 312 AVVHMRKLNHLNLSHPGVRPEVE 334
Db 305 LFLRHIRKL-----GOIPEGE 320

RESULT 5
US-09-272-796-5
; Sequence 5, Application US/09272796
; Patent No. 6207148
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/272,796  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/878,989  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0321 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 343 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PROSNOT06  
; CLONE: 827431  
; US-09-272-796-5

Query Match 39.4%; Score 990; DB 4; Length 343;  
Best Local Similarity 59.8%; Pred. No. 2.9e-79;  
Matches 193; Conservative 49; Mismatches 73; Indels 8; Gaps 2;  
  
QY 13 KQTTNIRKTFIFMEVLGSGAFSEVFLVKQRTGKLFALKCI-KKSPAFRDSLENEIAV 71  
Db 5 KKHTDISSVYEIRERLGSGAFSEVFLVLAQERGSAAHLVALKCIKKALRGKEALVENEIAV 64  
  
QY 72 LKIKIHENIVTDIEDYESTHYLYVMQLVSGGELFDRIILRGVYVTEKDALSLVIOQVLSAV 131  
Db 65 LRRISPNIVALEVDHESPSHLVLAELVTGGELFDRIIMERGSYTEKDALSHLVGQVLGAV 124  
  
QY 132 KYLHENGIVHRDLKPNLLYLTPTEENSKIMITDFGLSKMEQNGIMSTACGTPGYVAPEVL 191  
Db 125 SYLHSGIVHRDLKPNLLYATPEFESKIMVDSFGLSKIQAGNMLGTACGTPGYVAPELL 184  
  
QY 192 AQPKYKAVDCWISGIVITVILGCGPPFYEEETESKLFEKIKEGYEFESPFWDIDISESAK 251  
Db 185 EQPKYKAVDVWALGVISVILGCGPPFYDESDFELFSQILRASVYEDFPWDDISESGK 244  
  
QY 252 DFTCHLLEKDPNRYTCERKALSHPWIDGNLTALHRDIYPSVSIQIQNFAKSKWRQAFNA 311  
Db 245 DFTIHLERDLQKRTCCQALRDLMTFWDTGFGDRDILGFVSEQIRKNFAWTHWKRAFNA 304  
  
QY 312 AVVHMRKLMNLSHSPVRPEVE 334  
Db 305 LFLRHKL-----GQIPEGE 320

RESULT 6  
US-09-734-030-2  
; Sequence 2, Application US/09734030  
; Patent No. 6461846  
; GENERAL INFORMATION:  
; APPLICANT: BEASLEY, Ellen M.  
; APPLICANT: MERKLOV, Gennady  
; APPLICANT: KETCHUM, Karen A.  
; APPLICANT: WEI, Ming-Hui  
; APPLICANT: DIFRANCESCO, Valentina  
; APPLICANT: YAN, Chunhua  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL000612

; CURRENT APPLICATION NUMBER: US/09/734,030  
; CURRENT FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: 60/207,281  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 501  
; TYPE: PRT  
; ORGANISM: HUMAN  
; US-09-734-030-2  
  
Query Match 29.7%; Score 747.5; DB 4; Length 501;  
Best Local Similarity 35.3%; Pred. No. 1.2e-57;  
Matches 173; Conservative 86; Mismatches 170; Indels 61; Gaps 12;  
  
QY 15 QTTNIRKTFIFMEVLGSGAFSEVFLVKQRTGKLFALKCIKKSPAFRD-----SSLENEI 69  
Db 16 QPSEVTDYDLGQVIKTEEFCEIFRAKDKTKGLHTCKKFKQ-----RDGRKVRKAANEI 71  
  
QY 70 AVLKKIHENIVTLEDIYESTHYLYVMQLVSGGELFDRIILRGVYVTEKDALSLVIOQVLS 129  
Db 72 GILKMKVHPNQLQVDFVTRKEYFIFLELATGREVFDWILDQGYYSERDTSNVVRQVLE 131  
  
QY 130 AVKYLHENGIVHRDLKPNLLYLTPTEENSKIMITDFGLSKMEQNGIMSTACGTPGYVAPE 189  
Db 132 AVAYLSLKIIVHRNLKLENLVYNRLKNSKIVISDFHLAKLE-NGLIKEPCGTPPEYLAPE 190  
  
QY 190 VLAQPKYSKAVDCWISGIVITVILGCGPPFYEEETESKLFEKIKEGYEFESPE 241  
Db 191 VVGROYRGVDCWAGVIMILLSGNPPFYEEEDYENHDKNLFKILLAGYEDFSDP 250  
  
QY 242 FWDIDISESAKFICHLEKDPNRYTCERKALSHPWIDGNLTALHRDIYPSVSIQIQNFAK 301  
Db 251 YWDDISQAAKDLVTRLMEVEQDQRITAEAIKSHWISGNAASDKNIKDGVCQAIEKNFAR 310  
  
QY 302 SKWRQAFNAAVVHMRKLMNLSHSPVRPEVNRPPETOAS--ETSRP-----SSPE 352  
Db 311 AKWKAVRVTLMKRLR-----APEQSSATAAQSASATDTATPGAAGATATAA 359  
  
QY 353 ITITEAPVLDHSHVALPALTLQPCQHRRTAPGGRSLNCLVNGSL--HISSSLVPMHQGS 410  
Db 360 SGATSAPEGDAARAASDNVAPADRSATPATDG--SATPATDGSVTPATDGSITPATDGS 417  
  
QY 411 L-----AAGPCGCCSSCLNIGSKGKSYCSEPIPLKANKKQNFKSEVMV---PVRASG 461  
Db 418 VTPVTDGSRATP-----ATDGRATPATPEESTVPTTQSSAMLATKAAATPEPAMAQP 467  
  
QY 462 SSHCRAGOTG 471  
Db 468 DSTAPEGATG 477  
  
RESULT 7  
US-08-715-568A-1  
; Sequence 1, Application US/08715568A  
; Patent No. 5856463  
; GENERAL INFORMATION:  
; APPLICANT: Prydz, Hans Peter Blankenborg  
; APPLICANT: Brede, Gaute  
; TITLE OF INVENTION: PSKH-1 Ribozymes and Uses in Disease  
; TITLE OF INVENTION: Treatment  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lerner, David, Littenberg, Krumholz & Mentlik  
; STREET: 600 South Avenue West  
; CITY: Westfield  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07090-1497  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/715,568A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Foley, Shawn P.
; REFERENCE/DOCKET NUMBER: FORSK 3.0-002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-654-5000
; TELEFAX: 908-654-7866
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-715-568A-1

Query Match      25.9%; Score 650.5; DB 2; Length 424;
Best Local Similarity 40.2%; Pred. No. 3.5e-49;
Matches 131; Conservative 62; Mismatches 98; Indels 35; Gaps 4;

QY 27 EVLGSAFSEVFLVKQRLTGKLFALKCIKKSPAFRSDSSLENIATVLLKKIKHENIVTLEDI 86
Db 102 EIIGRGSFRRVVRVHRATROPYAIKMIETKYREGREVCESELVLRVRVHRHANIQLQVEV 161

QY 87 YESTHYLVMLQVSGGELFDRLILRGVYTEKDAISLVIOQVLSAVKYLHENGIVHRDLKP 146
Db 162 FETQERYVMVMEATGGELFDRIIAKGSFTRDATRVLQWLDGVRYLHALGITHRDLKP 221

QY 147 ENLLYLPENSKIMITDGLSKMEONG---IMSTAGTGPYVAPVLAQKPYSKAVDCW 203
Db 222 ENLLYHPGTDSKIIITDFGLASARKKGGDCLMKTTCTGPTPIAPEVLVRKPYTNSVDMW 281

QY 204 SIGVITYILLCGYPPFVEETESKLFKEIKGYEYEFSPFWDIDISESAKFCHLLEKDPN 263
Db 282 ALGVIAVLLSGTMTPEDDNRTLRQLRGKYSYSGEPWPSVSNLAKDFIDRLLTVDPG 341

QY 264 ERYTCEKALSHPWIDNTALHRDIYPSVLSIQKNEPAKSKWQAFNAAVVHMKRLMN 323
Db 342 ARMTALQALRHPWV-----VSM-----AASSSMKNLHRS 370

QY 324 LHSPGVRPEVENRPPETOASRSPS 349
Db 371 I-SQNLKRASSRCQSTKSAQSTRSS 395

RESULT 8
US-07-857-224B-24
; Sequence 24, Application US/07857224B
; Patent No. 5958784
; GENERAL INFORMATION:
; APPLICANT: Benner, Steven A.
; TITLE OF INVENTION: Predicting Folded Structures of Proteins
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven A. Benner
; STREET: Hadlaubstrasse 151
; CITY: Zurich
; STATE: none
; COUNTRY: Switzerland
; ZIP: (note: this is an international post code) CH-8092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/857,224B
; FILING DATE: 03/25/92
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA: none
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (International) 41 1 632 2830
; TELEFAX: (International) 41 1 262 2437
; TELEX: none
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 264
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; ORIGINAL SOURCE:
; ORGANISM: human
; FEATURE: Protein kinase; Table 8 Column 26
; PUBLICATION INFORMATION:
; AUTHORS:
; AUTHORS: Hanks, S. K.
; AUTHORS: Quinn, A. M.
; AUTHORS: Hunter, T.
; TITLE: The protein kinase family
; JOURNAL: Science
; VOLUME: 241
; PAGES: 42-52
; DATE: 1988
; US-07-857-224B-24

Query Match      24.9%; Score 625; DB 2; Length 264;
Best Local Similarity 46.9%; Pred. No. 3e-47;
Matches 119; Conservative 50; Mismatches 81; Indels 4; Gaps 2;

QY 28 VLGSGAFSEVFLVKQRLTGKLFALKCIKKSPAFRSDSSLENIATVLLKKIKHENIVTLEDIY 87
Db 8 LIGRGSFRRVVRVHRATROPYAIKMIETKYREGREVCESELVLRVRVHRHANIQLQVEVF 67

QY 88 ESTHYLVMLQVSGGELFDRLILRGVYTEKDAISLVIOQVLSAVKYLHENGIVHRDLKP 146
Db 68 ETQERYVMVMEATGGELFDRIIAKGSFTRDATRVLQWLDGVRYLHALGITHRDLKP 127

QY 147 ENLLYLPENSKIMITDGLSKMEONG---IMSTAGTGPYVAPVLAQKPYSKAVDCW 203
Db 128 ENLLYHPGTDSKIIITDFGLASARKKGGDCLMKTTCTGPTPIAPEVLVRKPYTNSVDMW 187

QY 204 SIGVITYILLCGYPPFVEETESKLFKEIKGYEYEFSPFWDIDISESAKFCHLLEKDPN 263
Db 188 ALGVIAVLLSGTMTPEDDNRTLRQLRGKYSYSGEPWPSVSNLAKDFIDRLLTVDPG 247

QY 264 ERYTCEKALSHPWI 277
Db 248 ARMTALQALRHPWV 261

RESULT 9
US-09-800-960-4
; Sequence 4, Application US/09800960
; Patent No. 6387677
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001158
; CURRENT APPLICATION NUMBER: US/09/800,960
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Human
; US-09-800-960-4

Query Match      23.6%; Score 592; DB 4; Length 556;
Best Local Similarity 33.2%; Pred. No. 7.5e-44;
```

Matches 149; Conservative 75; Mismatches 177; Indels 48; Gaps 9;

QY 17 TNIRKTFIMFVLGSGAFSEVFLVKQRLTGKLFALKCI--KKSPAFRSDSSLENEIAVLKK 74  
Db 8 TRFTDDYQLFELGKGFVVRRCVKKSTQDYAAKIINTKKLSARDHQKLEREARICRL 67  
QY 75 IKHENIVTLEDIYESTTHYLLVMQVSGGELFDRLIRGVYTEKDASLVIOQVLSAVKYL 134  
Db 68 LKHPNIVRLHDSISEGFHYLVDFDLVTGGELFEDIVAREYIYSEADASHCIHQILESVDH 127  
QY 135 HENGIVHRDLKPNLLYTPPENSKIMITDFGLS---KMEQNGIMSTACGTPGVVAPVL 191  
Db 128 HQHIVHRDLKPNLLKASKCKGAALKLADFLGAEVGEQQAQWFGFA-GTPGVLSPEVL 186  
QY 192 AOKPYSKAVDCWSIGVITVILCGVPPFYBETESKLFKIKEGYVEFPFDDISSEAK 251  
Db 187 RDPGKGVDIWACGVILYLLVGGPPFDEDOHKLIOIIRKAGAYDFPSPEDVTPPAK 246  
QY 252 DFICHLLEKDPNERTYCEKALSHPWIDGNTALHRDIYPSVSLQIQNF-AKSKWRQAFNA 310  
Db 247 NLINQMLTINPAKRITADQALKHPWCORSTVASMHRQETVECLRKFNARRKLGKALT 306  
QY 311 AAVVHH---MRKLNHLNHP-CVREVENRPPETQASSTRSPSPETITTEAPVLDHVA 366  
Db 307 TWLSRNFSAKSLNKKSDGGVKPQSNKK-----NSLV 340  
QY 367 LPALTQLPCQHGRRPTAPGGRSLNCLVNGSLHISS-----LVPMHQGSAAAGPGC 418  
Db 341 SPAQEPALQTAQMEPQTIVVHNATDGIKSTESCNTTTEDEDLKAAPLRTGNGSSVPBGR 400  
QY 419 CSSCLNIGSKGSSVCSEPTLLKANKKQ 447  
Db 401 SSRDRTPAGMQ---POPSLSSAMRKQ 426

RESULT 10  
US-07-857-224B-18  
; Sequence 18, Application US/07857224B  
; Patent No. 5958784  
; GENERAL INFORMATION:  
; APPLICANT: Benner, Steven A.  
; TITLE OF INVENTION: Predicting Folded Structures of Proteins  
; NUMBER OF SEQUENCES: 114  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Steven A. Benner  
; STREET: Hadlaubstrasse 151  
; CITY: Zurich  
; STATE: none  
; COUNTRY: Switzerland  
; ZIP: (note: this is an international post code) CH-8092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh 7.0  
; SOFTWARE: Microsoft Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/857,224B  
; FILING DATE: 03/25/92  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA: none  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (International) 41 1 632 2830  
; TELEFAX: (International) 41 1 262 2437  
; TELEX: none  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 264  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; DESCRIPTION: protein  
; ORIGINAL SOURCE:  
; ORGANISM: rat

; FEATURE: Protein kinase; Table 8 Column 20  
; PUBLICATION INFORMATION:  
; AUTHORS: Hanks, S. K.  
; AUTHORS: Quinn, A. M.  
; AUTHORS: Hunter, T.  
; TITLE: The protein kinase family  
; JOURNAL: Science  
; VOLUME: 241  
; PAGES: 42-52  
; DATE: 1988  
; US-07-857-224B-18

Query Match 23.4%; Score 587.5; DB 2; Length 264;  
Best Local Similarity 46.6%; Pred. No. 6.2e-44;  
Matches 123; Conservative 42; Mismatches 88; Indels 11; Gaps 4;

QY 27 EVLGSGAFSEVFLVKQRLTGKLFALKCI--KKSPAFRSDSSLENEIAVLKKIKHENIVTLE 84  
Db 7 EELGKGFVVRRCVKVLAGQEAARKIINTKKLSARDHQKLEREARICRLKKHPNIVRLH 66  
QY 85 DIVESTTHYLLVMQVSGGELFDRLIRGVYTEKDASLVIOQVLSAVKYLHENGIVHRDL 144  
Db 67 DSISEEGHLLIPDLVTGGELFEDIVAREYIYSEADASHCIQILEAVLHCHQMGVVRDL 126  
QY 145 KPNLLYLTPENSKIMITDFGLS---KMEQNGIMSTACGTPGVVAPVLAQPKYSKAVD 201  
Db 127 KPNLLASLAKGAALKLADFLGAEVGEQQAQWFGFA-GTPGVLSPEVLRKDPYKQPV 185  
QY 202 CWSIGVITYILLCGVPPFYBETESKLFKIKEGYVEFPFDDISSEAKDFICHLEKD 261  
Db 186 LWACGVILYLLVGGPPFDEDOHRLYQIQKAGAYDFPSPEDVTPPEAKDLINKMLTIN 245  
QY 262 PNERTYCEKALSHPWIDGNTALHR 285  
Db 246 PSKRITAAEALKHPWIS-----HR 264

RESULT 11  
US-09-529-093A-2  
; Sequence 2, Application US/09529093A  
; Patent No. 6413755  
; GENERAL INFORMATION:  
; APPLICANT: LUYTEN, Walter H.M.L.  
; APPLICANT: PARKER, Andrew E.  
; APPLICANT: MCGOWAN, Clare H.  
; APPLICANT: BLASINA, Alessandra  
; TITLE OF INVENTION: Human Checkpoint Kinase, hcdsl, Compositions and Methods  
; FILE REFERENCE: TSRI 649.0  
; CURRENT APPLICATION NUMBER: US/09/529, 093A  
; CURRENT FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: PCT/EP98/06981  
; PRIOR FILING DATE: 1998-10-21  
; PRIOR APPLICATION NUMBER: GB 9722320.0  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 543  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-529-093A-2

Query Match 23.0%; Score 578.5; DB 4; Length 543;  
Best Local Similarity 43.9%; Pred. No. 1.1e-42;  
Matches 126; Conservative 47; Mismatches 91; Indels 23; Gaps 7;

QY 7 DDCSSWKKQTINIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIK-----56  
Db 207 DDQSVYPKA---LRDEVIMSKTLGSGACGEVKLAFERKTKCKKVAIKIISKRKFAIGSARE 263  
QY 57 -SPAFRSDSSLENEIAVLKKIKHENIVTLEDIYESTTHYLLVMQVSGGELFDRLIRGVY 115

```
Db 264 ADPAL--NVETEIEILKLNHPICIIKKNFTDA-EDYIVVLEMEGGELFDKVVGNKRL 319
QY 116 TEKDASLVITQOVLSAVKYVHENGIVHRDLKPNENLYLTPENSKIMITDFGLSK-MEONG 174
Db 320 KEATCKLYFYQMLLAQVYLHENGIIHRDLKPNENLVSSQEDCLIKITDFGHSKILGETS 379
QY 175 IMSTAGCTPGYVAPEVLAQ--KPYSKAVDCWSIGVITVILLCGYPPPEYE-ETESKLEPK 230
Db 380 LMTLCTGTPTYLAPVLSVGTAGYNRAVDCWSLGVIFICLSGYPPFPSEHRTQVSLKDQ 439
QY 231 IKEYGYEFESPFWDIDISESAKDFICHLLEKDPNERTCEKALSHPMW 277
Db 440 ITSGKYNFPEVWAEYSERALDLVKLLVVDPKARETIEEALRHPWL 486

RESULT 12
US-09-800-960-2
; Sequence 2, Application US/098000960
; Patent No. 6387677
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01158
; CURRENT APPLICATION NUMBER: US/09/800,960
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Human
US-09-800-960-2

Query Match 22.7%; Score 571.5; DB 4; Length 565;
Best Local Similarity 34.4%; Pred. No. 4.9e-42;
Matches 142; Conservative 69; Mismatches 151; Indels 51; Gaps 10;

QY 17 TNRKTFIFMEVLGSGAFSEVLVKORLTGKLFALKCI--KKSPAFRDSSENEIAVLKK 74
Db 8 TRFTDDYQLFEELGKGAFSVVRCVKKTSTQYAAKIINTKKLSARDHQKLEREARICRL 67
QY 75 IKHENIVTLEDIYESTHYVLMOLVSGGELFDRLERGVYTEKDSLVITQOVLSAVKYL 134
Db 68 LKHPNIVRLHDSISEGPHYVFDLVGTGGELFDIVAREYYSEADASHCIHOILESVNHI 127
QY 135 HENGIVHRDLKPNENLYLTPENSKIMITDFGLS---KMEONGIMSTACGTPGYVAPEVL 191
Db 128 HOHDIVHRDLKPNENLLASKCKGAAVKLADFGLAIEVQGEQQAQWFGFA-GTPGYLSPEVL 186
QY 192 AOKPYSKAVDCWSIGVITVILLCGYPPPEYEETESKLFKEIKGYEYEFSPFWDIDISESAK 251
Db 187 RKDPYKGPVDIMACGVILVILVGYPPFWDDEQHKLYQQIKAGAYDFPSPENDVTYTPK 246
QY 252 DFICHLLEKDPNERTCEKALSHPMWIDGNTA---LHRDIYPS-----V 291
Db 247 NLINQMLTINPAKRITADQALKHPWCQSRSTVASMMHQETVECLRKFNARKKLKGAILT 306
QY 292 SLQIQKNFAKSKWRQ-----AFNAAAVVHMKRLHMLNHSQVPRPEVENRPETQASETS 346
Db 307 TMLVSRNFSVGRSSAPASPAASAGLAGAAKSLINKSDG---GVKKR-----KSSSV 359
QY 347 RPSSPETTI-----TEAPVLDSHVALPALTPQCGRRPTAPGGRS 388
Db 360 HLMPEQTIVVHNATDGIKSTES--CNTTDEDEDLKAAPLRTGTGSSVPEGRS 410

RESULT 13
US-07-857-224B-19
; Sequence 19, Application US/07857224B
; Patent No. 5938784
; GENERAL INFORMATION:
```

```
; APPLICANT: Benner, Steven A.
; TITLE OF INVENTION: Predicting Folded Structures of Proteins
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven A. Benner
; STREET: Hadlaubstrasse 151
; CITY: Zurich
; STATE: none
; COUNTRY: Switzerland
; ZIP: (note: this is an international post code) CH-8092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/857,224B
; FILING DATE: 03/25/92
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA: none
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (International) 41 1 632 2830
; TELEFAX: (International) 41 1 262 2437
; TELEX: none
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 264
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; ORIGINAL SOURCE:
; ORGANISM: rat
; FEATURE: Protein kinase; Table 8 Column 21
; PUBLICATION INFORMATION:
; AUTHORS: Hanks, S. K.
; AUTHORS: Quinn, A. M.
; AUTHORS: Hunter, T.
; TITLE: The protein kinase family
; JOURNAL: Science
; VOLUME: 241
; PAGES: 42-52
; DATE: 1988
; US-07-857-224B-19

Query Match 22.4%; Score 564; DB 2; Length 264;
Best Local Similarity 45.7%; Pred. No. 7.3e-42;
Matches 117; Conservative 42; Mismatches 91; Indels 6; Gaps 3;

QY 27 EVLGSAGFSEVLVKORLTGKLFALKCI--KKSPAFRDSSENEIAVLKKIKHENIVTLE 84
Db 7 EDIGKGAFSVVRCVKCTGHEYAAKIINTKKLSARDHQKLEREARICRLKHSNIVRLH 66
QY 85 DIYESTTHYVLMOLVSGGELFDRLERGVYTEKDSLVITQOVLSAVKYVHENGIVHRDL 144
Db 67 DSISEGPHYVLPDLVTGGELFDIVAREYYSEADASHCIQILEAVLHCHQMGVVRDL 126
QY 145 KPNENLYLTPENSKIMITDFGLSKMEONGIMST---ACGTPGYVAPEVLAQPYSKAVD 201
Db 127 KPNENLLASKLKGAAVKLADFGLA-IEVQGDQQAQWFGAGTGGYLSPEVLRKEAYGKPD 185
QY 202 CWSIGVITVILLCGYPPPEYEETESKLFKEIKGYEYEFSPFWDIDISESAKDFICHLLEK 261
Db 186 IWACGVILVILVGYPPFWDDEQHKLYQQIKAGAYDFPSPENDVTYTPKANLINQMLTIN 245
QY 262 PNERTCEKALSHPMW 277
Db 246 PAKRITAHEALKHPW 261

RESULT 14
US-07-951-715A-23
```

; Sequence 23, Application US/07951715A  
; Patent No. 5625136  
; GENERAL INFORMATION:  
; APPLICANT: Koziel, Michael G.  
; APPLICANT: Desai, Nalini M.  
; APPLICANT: Lewis, Kelly S.  
; APPLICANT: Kramer, Vance C.  
; APPLICANT: Warren, Gregory W.  
; APPLICANT: Evola, Stephen V.  
; APPLICANT: Crossland, Lyle D.  
; APPLICANT: Wright, Martha S.  
; APPLICANT: Merlin, Ellis J.  
; APPLICANT: Launis, Karen L.  
; APPLICANT: Rothstein, Steven J.  
; APPLICANT: Bowman, Cindy G.  
; APPLICANT: Dawson, John L.  
; APPLICANT: Dunder, Erik M.  
; APPLICANT: Pace, Gary M.  
; APPLICANT: Suttie, Janet L.  
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED  
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE  
; NUMBER OF SEQUENCES: 94  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CIBA-GEIGY Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10532  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/951,715A  
; FILING DATE: 25-SEP-1992  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/772,027  
; FILING DATE: 04-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spruill, W. Murray  
; REGISTRATION NUMBER: 32,943  
; REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (919)541-8615  
; TELEFAX: (919)541-8689  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 295 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEetical: NO  
; FEATURE:  
; NAME/KEY: protein  
; LOCATION: 1..295  
; OTHER INFORMATION: /note= "rat protein kinase II  
; OTHER INFORMATION: protein sequence as shown in Figure 32."  
; US-07-951-715A-23

Query Match 22.4%; Score 563.5; DB 1; Length 295;  
Best Local Similarity 41.8%; Pred. No. 9.5e-42;  
Matches 120; Conservative 56; Mismatches 104; Indels 7; Gaps 4;  
Qy 27 EVLGSGAFSEVFLVQRUTGKLFALKCI--KKSPAFRDSSELENIATVKKIKHNIVTLE 84  
Db 5 BELGKGFSEVVRRCVTKSTQEAAKIINTKKLSARDHQKLEAREARICRLKKHPNIVRLH 64  
Qy 85 DIYESTTHYVLVMOIVSGGLFDRITLGRGVYVTEKDAIVIOQVLSAVKYLHENGIVHRDL 144

Db 65 DSISEEGFHYLVFDLVTGGELFEDIVAREYYSEADASHCIHQILBSVNIHQHDIIVHRDL 124  
Qy 145 KPENLLYLTPEENSKIMITDFGLS---KMEQNGIMSTACGTFCYVAPEVLAQKPYSKAYD 201  
Db 125 KPENLLASKCKGAAVKLADFLAIEVQEQQAQWFGFA-GTPGYLSPEVLRKDPYKGPVD 183  
Qy 202 CWSIGVITVILLCGYPFYEETESKLFKEIKEGYEFESPFWDIDISESAKDFICHLLKED 261  
Db 184 IWACGVILYLLVGYPPFWEDEQHKLYQOIKAGAYDFPSPEDWTVTPEAKNLINQMLTIN 243  
Qy 262 PNERVTCERKALSHPWIDGNTALHRIYPSVSLQIQKNF-AKSKWROA 307  
Db 244 PAKRITADQALKHPWVCQRSTVASMHRQETVECLRKFNARKLKGA 290  
RESULT 15  
US-08-459-448A-23  
; Sequence 23, Application US/08459448A  
; Patent No. 5859336  
; GENERAL INFORMATION:  
; APPLICANT: Koziel, Michael G.  
; APPLICANT: Desai, Nalini M.  
; APPLICANT: Lewis, Kelly S.  
; APPLICANT: Kramer, Vance C.  
; APPLICANT: Warren, Gregory W.  
; APPLICANT: Evola, Stephen V.  
; APPLICANT: Crossland, Lyle D.  
; APPLICANT: Wright, Martha S.  
; APPLICANT: Merlin, Ellis J.  
; APPLICANT: Launis, Karen L.  
; APPLICANT: Rothstein, Steven J.  
; APPLICANT: Bowman, Cindy G.  
; APPLICANT: Dawson, John L.  
; APPLICANT: Dunder, Erik M.  
; APPLICANT: Pace, Gary M.  
; APPLICANT: Suttie, Janet L.  
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED  
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE  
; NUMBER OF SEQUENCES: 94  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5859336artis Corporation  
; STREET: Patent & Trademark Dept., 520 White Plains  
; STREET: Rd., POB 2005  
; CITY: Tarrytown  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10591-9005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/459,448A  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/951,715  
; FILING DATE: 25-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/772,027  
; FILING DATE: 04-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pace, Gary M.  
; REGISTRATION NUMBER: 40403  
; REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (919)541-8582  
; TELEFAX: (919)541-8689  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 295 amino acids  
; TYPE: amino acid

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..295
; OTHER INFORMATION: /note= "rat protein kinase II
; US-08-459-448A-23

Query Match      22.4%; Score 563.5; DB 2; Length 295;
Best Local Similarity 41.8%; Pred. No. 9.5e-42;
Matches 120; Conservative 56; Mismatches 104; Indels 7; Gaps 4;

QY 27 EVLGSGAFSEVFLVKQRLTGKLFALKCI--KKSPAFPRDSSLENEIAVLKKIKHENIVTLE 84
Db 5 EELGKGAFSVVRCVKKTSTQEYAAKIINTKKLSARDHQKLEREARICRLKHPNIVRLH 64
QY 85 DIYESTHYLYVMQLYSGGELFDRILERGVYTEKDASLVIOQVLSAVKYLHENGIVHRDL 144
Db 65 DSISEEGFHYLVFDLVTGGELFEDIVAREYIYSEADASHCIHQLESVNHQHDIIVHRDL 124
QY 145 KPNLLVLTPEENSKIMITDFGLS--KWEONGIMSTAGTGCYVAPEVLAOKPYSKAYD 201
Db 125 KPNLLASCKGAAYKLADFGIAIEVQEQQAWFGFA-GTPGYLSPVLRKDPYGPYD 183
QY 202 CWSIGVITYILLCGYPFPYEETESKLFKEIKGYEFESPFWDIDISEAKDFICHLLLEKD 261
Db 184 IWACGVILYILLVGYPPFWEQDKLVYQKIKAGAYDFPSPWMDTPTPEAKNLINQMLTIN 243
QY 262 PNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIQKNF-AKSKWROA 307
Db 244 PAKRITADQALKHPWVCQRSTVASMHRQETVECLRKFNARRKLKGA 290
```

Search completed: March 14, 2003, 14:10:30  
Job time : 19 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:06:52 ; Search time 32 Seconds  
(without alignments)  
3064.951 Million cell updates/sec

Title: US-09-960-643-2  
Perfect score: 2513  
Sequence: 1 MKREDDCSWKKQTNR.....VKASGSHCRAGTGVCLIM 476

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.21.\*

1: sp.archaea.\*  
2: sp.bacteria.\*  
3: sp.fungi.\*  
4: sp.human.\*  
5: sp.invertebrate.\*  
6: sp.mammal.\*  
7: sp.mhc.\*  
8: sp.organella.\*  
9: sp.phage.\*  
10: sp.plant.\*  
11: sp.rodent.\*  
12: sp.virus.\*  
13: sp.invertebrate.\*  
14: sp.unclassified.\*  
15: sp.virus.\*  
16: sp.bacteriaph.\*  
17: sp.archaea.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2513	100.0	481	4	Q9V3J7	Q9V3J7 homo sapien
2	2509	99.8	476	4	Q96NX5	Q96NX5 homo sapien
3	2362	94.0	460	4	Q95523	Q95523 homo sapien
4	2327.5	92.6	477	11	Q91VB2	Q91VB2 mus musculus
5	1612	64.1	309	11	Q08763	Q08763 rattus norv
6	1246	49.6	357	4	Q9HD31	Q9HD31 homo sapien
7	1155.5	46.0	374	11	Q91VS8	Q91VS8 mus musculus
8	1070	42.6	343	11	Q70150	Q70150 rattus norv
9	1068	42.5	342	11	Q08767	Q08767 rattus norv
10	1064	42.3	343	11	Q9QYK9	Q9QYK9 mus musculus
11	1047	41.7	343	4	Q9NNY2	Q9NNY2 homo sapien
12	1008	40.1	348	5	Q9TXJ0	Q9TXJ0 caenorhabdi
13	1004	40.0	348	5	Q9UAH6	Q9UAH6 caenorhabdi
14	957.5	38.1	405	5	Q9V3I4	Q9V3I4 drosophila
15	747.5	29.7	501	4	Q8WTT8	Q8WTT8 homo sapien
16	747.5	29.7	501	4	Q9BQC9	Q9BQC9 homo sapien

17	739	29.4	512	11	Q8VD20	Q8VD20 mus musculus
18	736	29.3	504	11	Q63092	Q63092 rattus norv
19	703	28.0	421	13	Q9YGM4	Q9YGM4 fugu rubrip
20	692.5	27.6	502	11	Q63892	Q63892 rattus sp.
21	674	26.8	385	13	Q98T22	Q98T22 xenopus lae
22	651.5	25.9	349	5	Q8SSR6	Q8SSR6 dictyosteli
23	643.5	25.6	424	11	Q91YA2	Q91YA2 mus musculus
24	643	25.6	373	3	Q9HF33	Q9HF33 arthrobotry
25	635	25.3	420	3	Q42627	Q42627 collettotric
26	631.5	25.1	415	3	Q96UI7	Q96UI7 neosporea
27	628.5	25.0	473	4	Q9HQ05	Q9HQ05 homo sapien
28	623	24.8	404	3	Q9Y899	Q9Y899 emeritella
29	621	24.7	513	5	Q15865	Q15865 plasmodium
30	615.5	24.5	533	5	Q9NH59	Q9NH59 caenorhabdi
31	615.5	24.5	571	5	Q9NH58	Q9NH58 caenorhabdi
32	610	24.3	720	5	Q62305	Q62305 caenorhabdi
33	608	24.2	518	5	Q9NH57	Q9NH57 caenorhabdi
34	606.5	24.1	385	4	Q96QS6	Q96QS6 homo sapien
35	605.5	24.1	433	11	Q9JLM6	Q9JLM6 mus musculus
36	604.5	24.1	589	11	Q63094	Q63094 rattus norv
37	603	24.0	350	5	Q9NG91	Q9NG91 caenorhabdi
38	603	24.0	482	5	Q9U6Q0	Q9U6Q0 caenorhabdi
39	603	24.0	520	5	Q9NH60	Q9NH60 caenorhabdi
40	602	24.0	478	13	Q9YHB8	Q9YHB8 gallus gall
41	602	24.0	527	4	Q9Y2H4	Q9Y2H4 homo sapien
42	600	23.9	530	5	Q00168	Q00168 drosophila
43	599.5	23.9	493	5	Q9V495	Q9V495 drosophila
44	598.5	23.8	489	13	Q93559	Q93559 gallus gall
45	598.5	23.8	499	6	Q95266	Q95266 sus scrofa

## ALIGNMENTS

RESULT 1

Q9Y3J7 PRELIMINARY; PRT; 481 AA.  
ID Q9Y3J7  
AC Q9Y3J7;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Hypothetical 53.6 kDa protein (fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Rhodes S.;  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AL049688; CAB41259.1; .  
DR HSSP; Q63450; 1A06.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR001230; Prenyl\_site.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD0000001; Euk\_pkinase; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR PROSITE; PS00294; PRENYLATION; UNKNOWN\_1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_SP; 1.  
KW Hypothetical protein; Serine/threonine-protein kinase.  
FT NON\_TER 1  
SQ SEQUENCE 481 AA; 53588 MW; 3C5A4F9B6765333E CRC64;

Query Match 100.0%; Score 2513; DB 4; Length 481;  
Best Local Similarity 100.0%; Pred. No. 2.3e-190;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MGRKEEDDCSSWKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKKSPAF 60
Db 6 MGRKEEDDCSSWKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKKSPAF 65
QY 61 RDSSLENEIAVLKIKHENIVTLEDIYESTHYILVMQLVSGGELFDRILRGVYTEKDA 120
Db 66 RDSSLENEIAVLKIKHENIVTLEDIYESTHYILVMQLVSGGELFDRILRGVYTEKDA 125
QY 121 SLVIOQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKIMITDGLSKMEQNGIMSTAC 180
Db 126 SLVIOQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKIMITDGLSKMEQNGIMSTAC 185
QY 181 GTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPFPYEETESKLFKEIKEGYEFES 240
Db 186 GTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPFPYEETESKLFKEIKEGYEFES 245
QY 241 PFWDIDISESAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIQKNFA 300
Db 246 PFWDIDISESAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIQKNFA 305
QY 301 KSKWROAFNAAAVVHMRKLNHLSHSPGVPRPEVENRPPETQASSETSRRSPSEITITEAPV 360
Db 306 KSKWROAFNAAAVVHMRKLNHLSHSPGVPRPEVENRPPETQASSETSRRSPSEITITEAPV 365
QY 361 LDHSVALPALTLQPCQHGRRPTAPGGRSLNCLVNGSLHISSSLVPMHOGSLAAGPCGCCS 420
Db 366 LDHSVALPALTLQPCQHGRRPTAPGGRSLNCLVNGSLHISSSLVPMHOGSLAAGPCGCCS 425
QY 421 SCNLGSKGKSSCYSEPTLLKANKKQNFSEVMPVKASGSSHCRAGOTGVCLIM 476
Db 426 SCNLGSKGKSSCYSEPTLLKANKKQNFSEVMPVKASGSSHCRAGOTGVCLIM 481

RESULT 2
Q96NX5 PRELIMINARY; PRT; 476 AA.
ID Q96NX5;
AC Q96NX5;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Calcium/calmodulin-dependent protein kinase I gamma.
GN CAMK1G.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
.OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20113118; PubMed=10645953;
RA Schutte B.C., Bjork B.C., Coppage K.B., Malik M.I., Gregory S.G.,
RA Scott D.J., Brentzell L.M., Watanabe Y., Dixon M.J., Murray J.C.;
RT "A preliminary gene map for the Van der Woude syndrome critical region
RT derived from 900 kb of genomic sequence at 1q32-q41."
RL Genome Res. 10:81-94(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Bjork B.C., Watanabe Y., Murray J.C., Schutte B.C.;
RT "Characterization of the human ortholog of rat Cam Kinase I gamma
RT (CamKig) at 1q32-q41."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF428261; AAL28100.1;
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR001230; Prenyl_site.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00219; TyrKc; 1.
DR PROSITE: PS00294; PHENYLATION; UNKNOWN_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Kinase; Transferase.
```

```
SQ SEQUENCE 476 AA; 53128 MW; DD0AEAA016E7506E CRC64;
Query Match 99.8%; Score 2509; DB 4; Length 476;
Best Local Similarity 99.6%; Pred. No. 4.8e-190;
Matches 474; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRKEEDDCSSWKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKKSPAF 60
Db 1 MGRKEEDDCSSWKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKKSPAF 60
QY 61 RDSSLENEIAVLKIKHENIVTLEDIYESTHYILVMQLVSGGELFDRILRGVYTEKDA 120
Db 61 RDSSLENEIAVLKIKHENIVTLEDIYESTHYILVMQLVSGGELFDRILRGVYTEKDA 120
QY 121 SLVIOQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKIMITDGLSKMEQNGIMSTAC 180
Db 121 SLVIOQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKIMITDGLSKMEQNGIMSTAC 180
QY 181 GTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPFPYEETESKLFKEIKEGYEFES 240
Db 181 GTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPFPYEETESKLFKEIKEGYEFES 240
QY 241 PFWDIDISESAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIQKNFA 300
Db 241 PFWDIDISESAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIQKNFA 300
QY 301 KSKWROAFNAAAVVHMRKLNHLSHSPGVPRPEVENRPPETQASSETSRRSPSEITITEAPV 360
Db 301 KSKWROAFNAAAVVHMRKLNHLSHSPGVPRPEVENRPPETQASSETSRRSPSEITITEAPV 360
QY 361 LDHSVALPALTLQPCQHGRRPTAPGGRSLNCLVNGSLHISSSLVPMHOGSLAAGPCGCCS 420
Db 361 LDHSVALPALTLQPCQHGRRPTAPGGRSLNCLVNGSLHISSSLVPMHOGSLAAGPCGCCS 420
QY 421 SCNLGSKGKSSCYSEPTLLKANKKQNFSEVMPVKASGSSHCRAGOTGVCLIM 476
Db 421 SCNLGSKGKSSCYSEPTLLKANKKQNFSEVMPVKASGSSHCRAGOTGVCLIM 476

RESULT 3
Q95523 PRELIMINARY; PRT; 460 AA.
ID Q95523;
AC Q95523;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE DJ272116.1 (Rat Ca2+/calmodulin dependent protein kinase like
DE protein).
DE DJ272116.1.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
.OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AL023754; CAAL9296.1; -.
DR HSSP: Q63450; 1A06.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 460 AA; 51486 MW; A16971DC50344496 CRC64;
Query Match 94.0%; Score 2362; DB 4; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.9e-178;
```



```
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase.1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00118; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00101; PROTEIN_KINASE_ST; 1.
DR Serine/threonine-protein kinase.
KW NON_TER 309 309
FT
SQ SEQUENCE 309 AA; 35623 MW; 9162487561CF44E7 CRC64;

Query Match 64.1%; Score 1612; DB 11; Length 309;
Best Local Similarity 98.4%; Pred. No. 2.2e-119;
Matches 304; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGRKEEDCCSWKKQTNIIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKKSPAF 60
Db 1 MGRKEEDCCSWKKQTNIIRKTFIFMEVLGSGAFSEVFLVKQRTGKLFALKCIKKSPAF 60
Qy 61 ROSSLENEIATLVKIKHENIVTLEDIYESTHYLYVMQVLSGGELFDRILRGVYTEKDA 120
Db 61 ROSSLENEIATLVKIKHENIVTLEDIYESTHYLYVMQVLSGGELFDRILRGVYTEKDA 120
Qy 121 SLVIOQVLSAVKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMEQNGIMSTAC 180
Db 121 SLVIOQVLSAVKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMEQNGVMSTAC 180
Qy 181 GTPGVYVPEVLAQKPYSKAVDCWSIGVITYILLCGYPFPFYETESKLFKEIKEGYEFES 240
Db 181 GTPGVYVPEVLAQKPYSKAVDCWSIGVITYILLCGYPFPFYETESKLFKEIKEGYEFES 240
Qy 241 PFWDIDISESAKDFICHLLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIKNFA 300
Db 241 PFWDIDISESAKDFICHLLEKDPNERYTCEKALRHPWIDGNTALHRDIYPSVSLQIKNFA 300
Qy 301 KSKWROAFN 309
Db 301 KSKWROAFN 309

RESULT 6
Q9HD31 PRELIMINARY; PRT; 357 AA.
AC Q9HD31;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE CamK1-like protein kinase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Verploegen S., Koenderman L., Coffey P.J.;
RT "Identification and characterization of CKLIK: a novel granulocyte
RT Ca2+/calmodulin-dependent kinase."
RL Blood 0:0-0(2000).
DR EMBL: AF286366; AAG00534.1; -.
DR HSP: Q63450; IAO6
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR SMART: SM00219; TyrKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00118; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00101; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 357 AA; 40189 MW; 1FA184EEFA976FB4 CRC64;
```

```
Query Match 49.6%; Score 1246; DB 4; Length 357;
Best Local Similarity 73.1%; Pred. No. 2.3e-90;
Matches 237; Conservative 40; Mismatches 45; Indels 2; Gaps 2;

Qy 1 MGRKEEDCCSWKKQTNIIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPA 59
Db 1 MARENGESSSSWKKQAEDIKKIFEFKTLGTGAFSEVFLAEAEATGKLFVAVKIPKALK 60
Qy 60 FROSSLENEIATLVKIKHENIVTLEDIYESTHYLYVMQVLSGGELFDRILRGVYTEK 119
Db 61 GKESSENEIATLVKIKHENIVALEDIYESPNIHLIYVMQVLSGGELFDRIVEKGFTEK 120
Qy 120 ASLVIOQVLSAVKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMEQNG-IMST 178
Db 121 ASTLIROVLDVAVYVLRMGIVHRDLKPENLLYYSQDESKIMISDFGLSKMEKGDVNST 180
Qy 179 ACGTPGVYVPEVLAQKPYSKAVDCWSIGVITYILLCGYPFPFYETESKLFKEIKEGYEF 238
Db 181 ACGTPGVYVPEVLAQKPYSKAVDCWSIGVITYILLCGYPFPFYDENSKLFEQILKAEYEF 240
Qy 239 ESPWDIDISESAKDFICHLLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIKN 298
Db 241 DSPWDDIDISESAKDFINLMEKDPNKRYTCEQARHPWIAAGDTALNKNIHESVAQIRKN 300
Qy 299 FAKSKWROAFNAAAVVHMRKLHM 322
Db 301 FAKSKWROAFNATAVVRHMRKLHL 324

RESULT 7
Q91YS8 PRELIMINARY; PRT; 374 AA.
AC Q91YS8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Similar to calcium/calmodulin-dependent protein kinase I.
GN AI505105.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC014825; AAH14825.1; -.
DR MGI: 2141437; AI505105.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS00118; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00101; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 374 AA; 41624 MW; 37889CDA717D3AB2 CRC64;

Query Match 46.0%; Score 1155.5; DB 11; Length 374;
Best Local Similarity 69.6%; Pred. No. 3.5e-83;
Matches 222; Conservative 43; Mismatches 51; Indels 3; Gaps 3;

Qy 12 WKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPAPRDSLENEIA 70
Db 10 W-KQAEIRDIYDFRDVLGTGAFSEVILAEKRTQKLVAKIACAKALEGEGSMENEIA 68
Qy 71 VLKTKIKHENIVTLEDIYESTHYLYVMQVLSGGELFDRILRGVYTEKDAVLQOVLVA 130
Db 69 VLHKIKHPNIVALDDIYEGGHLIYLMQVLSGGELFDRIVEKGFTEKDAKRLIFQVLDA 128
Qy 131 VKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMEQNG-IMSTACGTFGYVAPE 189
Db 131 VKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMEQNG-IMSTACGTFGYVAPE 189
```

```
Db 129 VKYLDLGIHVHDLKPNENLYYSLDSDSKIMISDFGLSKMEDPGSVLSTAGTGGYVAPE 188
QY 190 VLAQKPYSKAVDCWSIGVTYILLGCGPPFYETESKLFKEIKEGYEFESPFWDIDISES 249
Db 189 VLAQKPYSKAVDCWSIGVTYILLGCGPPFYDENDAKLFEQLTKAEYEFDSFYWDIDISDS 248
QY 250 AKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDLYPSVSLQIQKNAFKSKWROAFN 309
Db 249 AKDFIRHLEKDPNERYTCEKALSHPWIDGNTALHRDLYPSVSLQIQKNAFKSKWROAFN 308
QY 310 AAQVVRHMKRLHMLNLSHSPG 328
Db 309 ATAVVRHMKRLQLGTSQEG 327

RESULT 8
ID 070150 PRELIMINARY; PRT; 343 AA.
AC 070150;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Ca2+/calmodulin-dependent protein kinase I beta 2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=98070455; PubMed=9405489;
RA Naito Y., Watanabe Y., Yokokura H., Sugita R., Nishio M., Hidaka H.;
RT "Isoform-specific activation and structure diversity of calmodulin
kinase I.";
RL J. Biol. Chem. 272:32704-32708(1997).
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB004267; BAA28263.1; -.
DR HSSP; Q63450; 1A06.
DR InterPro; IPR000719; Euk_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 343 AA; 38438 MW; 994D451D809A9D80 CRC64;

Query Match 42.6%; Score 1070; DB 11; Length 343;
Best Local Similarity 59.5%; Pred. No. 1.8e-76;
Matches 207; Conservative 55; Mismatches 74; Indels 12; Gaps 3;

QY 13 KKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRTGKLFALKCI-KKSPAFRDSLENEIATV 71
Db 5 KKQTEDISSVYEIREKLGSGAFSEVLAQERGAHLVALKCIKKALRGKEALVENEIATV 64
QY 72 LKIKIHENIVTLDIYESTHYLVNQLVSGGELFDRILERGVYTEKDALSLVIOQVLSAV 131
Db 65 LRRISHPNIVALEDVHESPESHLYLAMELVGTGGELFDRIMERGVYTEKDAHLVGQVLGN 124
QY 132 KYLHENGIVHRDLKPNENLYLTPPEENSKIMITDFGLSKMEQNGIMSTAGTGGYVAPEVL 191
Db 125 SYLHSLGIHVHRDLKPNENLYATPFEDSKINVSDFGLSKIQAGNMLGTAGTGGYVAPELL 184
QY 192 AQPKYKAVDCWSIGVTYILLGCGPPFYETESKLFKEIKEGYEFESPFWDIDISESAK 251
Db 185 EQPKYKAVDVWALGVISITLGCYPPFYDESDFGLSKIQAGNMLGTAGTGGYVAPELL 244
QY 252 DFICHLLEKDPNERYTCEKALSHPWIDGNTALHRDLYPSVSLQIQKNAFKSKWROAFNA 311
Db 245 DFIRHLLERDPQKRFTCQALQHLWISGDAALDRDILGVSVEQIQKNFARTHWKRAFNA 304

Query Match 42.5%; Score 1068; DB 11; Length 342;
Best Local Similarity 60.1%; Pred. No. 2.6e-76;
Matches 206; Conservative 53; Mismatches 70; Indels 14; Gaps 3;

QY 13 KKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRTGKLFALKCI-KKSPAFRDSLENEIATV 71
Db 5 KKQTEDISSVYEIREKLGSGAFSEVLAQERGAHLVALKCIKKALRGKEALVENEIATV 64
QY 72 LKIKIHENIVTLDIYESTHYLVNQLVSGGELFDRILERGVYTEKDALSLVIOQVLSAV 131
Db 65 LRRISHPNIVALEDVHESPESHLYLAMELVGTGGELFDRIMERGVYTEKDAHLVGQVLGN 124
QY 132 KYLHENGIVHRDLKPNENLYLTPPEENSKIMITDFGLSKMEQNGIMSTAGTGGYVAPEVL 191
Db 125 SYLHSLGIHVHRDLKPNENLYATPFEDSKINVSDFGLSKIQAGNMLGTAGTGGYVAPELL 184
QY 192 AQPKYKAVDCWSIGVTYILLGCGPPFYETESKLFKEIKEGYEFESPFWDIDISESAK 251
Db 185 EQPKYKAVDVWALGVISITLGCYPPFYDESDFGLSKIQAGNMLGTAGTGGYVAPELL 244
QY 252 DFICHLLEKDPNERYTCEKALSHPWIDGNTALHRDLYPSVSLQIQKNAFKSKWROAFNA 311
Db 245 DFIRHLLERDPQKRFTCQALQHLWISGDAALDRDILGVSVEQIQKNFARTHWKRAFNA 304

Query Match 42.5%; Score 1068; DB 11; Length 342;
Best Local Similarity 60.1%; Pred. No. 2.6e-76;
Matches 206; Conservative 53; Mismatches 70; Indels 14; Gaps 3;

QY 13 KKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRTGKLFALKCI-KKSPAFRDSLENEIATV 71
Db 5 KKQTEDISSVYEIREKLGSGAFSEVLAQERGAHLVALKCIKKALRGKEALVENEIATV 64
QY 72 LKIKIHENIVTLDIYESTHYLVNQLVSGGELFDRILERGVYTEKDALSLVIOQVLSAV 131
Db 65 LRRISHPNIVALEDVHESPESHLYLAMELVGTGGELFDRIMERGVYTEKDAHLVGQVLGN 124
QY 132 KYLHENGIVHRDLKPNENLYLTPPEENSKIMITDFGLSKMEQNGIMSTAGTGGYVAPEVL 191
Db 125 SYLHSLGIHVHRDLKPNENLYATPFEDSKINVSDFGLSKIQAGNMLGTAGTGGYVAPELL 184
QY 192 AQPKYKAVDCWSIGVTYILLGCGPPFYETESKLFKEIKEGYEFESPFWDIDISESAK 251
Db 185 EQPKYKAVDVWALGVISITLGCYPPFYDESDFGLSKIQAGNMLGTAGTGGYVAPELL 244
QY 252 DFICHLLEKDPNERYTCEKALSHPWIDGNTALHRDLYPSVSLQIQKNAFKSKWROAFNA 311
Db 245 DFIRHLLERDPQKRFTCQALQHLWISGDAALDRDILGVSVEQIQKNFARTHWKRAFNA 304

RESULT 10
ID 090YK9 PRELIMINARY; PRT; 343 AA.
AC 090YK9;
QY 090YK9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
```

```
QY 312 AVVHMRKLMNLSHSPGVREVENRPPETQASSETSRPSPSPSEITITAP 359
Db 305 SFLRHIRKL-----GQSPEGE-----EASROGMTRHSPGLGTQOSP 341

RESULT 9
ID 008767 PRELIMINARY; PRT; 342 AA.
AC 008767;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Protein kinase.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=97228532; PubMed=9074610;
RA Yokokura H., Terada O., Naito Y., Hidaka H.;
RT "Isolation and comparison of rat cDNAs encoding Ca2+/calmodulin-
dependent protein kinase I isoforms.";
RL Biochim. Biophys. Acta 1338:8-12(1997).
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; D86556; BAA19879.1; -.
DR HSSP; Q63450; 1A06.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 342 AA; 38463 MW; 546C464913E784A2 CRC64;

Query Match 42.5%; Score 1068; DB 11; Length 342;
Best Local Similarity 60.1%; Pred. No. 2.6e-76;
Matches 206; Conservative 53; Mismatches 70; Indels 14; Gaps 3;

QY 13 KKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRTGKLFALKCI-KKSPAFRDSLENEIATV 71
Db 5 KKQTEDISSVYEIREKLGSGAFSEVLAQERGAHLVALKCIKKALRGKEALVENEIATV 64
QY 72 LKIKIHENIVTLDIYESTHYLVNQLVSGGELFDRILERGVYTEKDALSLVIOQVLSAV 131
Db 65 LRRISHPNIVALEDVHESPESHLYLAMELVGTGGELFDRIMERGVYTEKDAHLVGQVLGN 124
QY 132 KYLHENGIVHRDLKPNENLYLTPPEENSKIMITDFGLSKMEQNGIMSTAGTGGYVAPEVL 191
Db 125 SYLHSLGIHVHRDLKPNENLYATPFEDSKINVSDFGLSKIQAGNMLGTAGTGGYVAPELL 184
QY 192 AQPKYKAVDCWSIGVTYILLGCGPPFYETESKLFKEIKEGYEFESPFWDIDISESAK 251
Db 185 EQPKYKAVDVWALGVISITLGCYPPFYDESDFGLSKIQAGNMLGTAGTGGYVAPELL 244
QY 252 DFICHLLEKDPNERYTCEKALSHPWIDGNTALHRDLYPSVSLQIQKNAFKSKWROAFNA 311
Db 245 DFIRHLLERDPQKRFTCQALQHLWISGDAALDRDILGVSVEQIQKNFARTHWKRAFNA 304

Query Match 42.5%; Score 1068; DB 11; Length 342;
Best Local Similarity 60.1%; Pred. No. 2.6e-76;
Matches 206; Conservative 53; Mismatches 70; Indels 14; Gaps 3;

QY 13 KKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRTGKLFALKCI-KKSPAFRDSLENEIATV 71
Db 5 KKQTEDISSVYEIREKLGSGAFSEVLAQERGAHLVALKCIKKALRGKEALVENEIATV 64
QY 72 LKIKIHENIVTLDIYESTHYLVNQLVSGGELFDRILERGVYTEKDALSLVIOQVLSAV 131
Db 65 LRRISHPNIVALEDVHESPESHLYLAMELVGTGGELFDRIMERGVYTEKDAHLVGQVLGN 124
QY 132 KYLHENGIVHRDLKPNENLYLTPPEENSKIMITDFGLSKMEQNGIMSTAGTGGYVAPEVL 191
Db 125 SYLHSLGIHVHRDLKPNENLYATPFEDSKINVSDFGLSKIQAGNMLGTAGTGGYVAPELL 184
QY 192 AQPKYKAVDCWSIGVTYILLGCGPPFYETESKLFKEIKEGYEFESPFWDIDISESAK 251
Db 185 EQPKYKAVDVWALGVISITLGCYPPFYDESDFGLSKIQAGNMLGTAGTGGYVAPELL 244
QY 252 DFICHLLEKDPNERYTCEKALSHPWIDGNTALHRDLYPSVSLQIQKNAFKSKWROAFNA 311
Db 245 DFIRHLLERDPQKRFTCQALQHLWISGDAALDRDILGVSVEQIQKNFARTHWKRAFNA 304

Query Match 42.5%; Score 1068; DB 11; Length 342;
Best Local Similarity 60.1%; Pred. No. 2.6e-76;
Matches 206; Conservative 53; Mismatches 70; Indels 14; Gaps 3;

QY 13 KKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRTGKLFALKCI-KKSPAFRDSLENEIATV 71
Db 5 KKQTEDISSVYEIREKLGSGAFSEVLAQERGAHLVALKCIKKALRGKEALVENEIATV 64
QY 72 LKIKIHENIVTLDIYESTHYLVNQLVSGGELFDRILERGVYTEKDALSLVIOQVLSAV 131
Db 65 LRRISHPNIVALEDVHESPESHLYLAMELVGTGGELFDRIMERGVYTEKDAHLVGQVLGN 124
QY 132 KYLHENGIVHRDLKPNENLYLTPPEENSKIMITDFGLSKMEQNGIMSTAGTGGYVAPEVL 191
Db 125 SYLHSLGIHVHRDLKPNENLYATPFEDSKINVSDFGLSKIQAGNMLGTAGTGGYVAPELL 184
QY 192 AQPKYKAVDCWSIGVTYILLGCGPPFYETESKLFKEIKEGYEFESPFWDIDISESAK 251
Db 185 EQPKYKAVDVWALGVISITLGCYPPFYDESDFGLSKIQAGNMLGTAGTGGYVAPELL 244
QY 252 DFICHLLEKDPNERYTCEKALSHPWIDGNTALHRDLYPSVSLQIQKNAFKSKWROAFNA 311
Db 245 DFIRHLLERDPQKRFTCQALQHLWISGDAALDRDILGVSVEQIQKNFARTHWKRAFNA 304

RESULT 10
ID 090YK9 PRELIMINARY; PRT; 343 AA.
AC 090YK9;
QY 090YK9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
```

```
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE MCAMK1-BETA2 protein (pregnancy upregulated NONBIQUITOUS
DE CA2+/calmodulin-dependent kinase PNCK).
GN PNCK OR CAMK1B OR MCAMK1-BETA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Ueda T., Minami Y.;
RT "mcamk1-beta2.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=20139438; PubMed=10673339;
RA Gardner H.P., Rajan J.V., Ha S.I., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A., Marquis S.T., Chodosh L.A.;
RT "Cloning, Characterization, and Chromosomal Localization of Pnck, a
RT Ca2+/Calmodulin-Dependent Protein Kinase.";
RL Genomics 63:279-288(2000).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB023027; BAA87926.1; -.
DR EMBL; AF181984; AAF29157.1; -.
DR HSSP; Q63450; 1A06.
DR MGD; MGI:1347357; Pnck.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 343 AA; 38519 MW; 1B4A28D367A936E CRC64;

Query Match 42.3%; Score 1064; DB 11; Length 343;
Best Local Similarity 59.2%; Pred. No. 5.4e-76;
Matches 206; Conservative 55; Mismatches 75; Indels 12; Gaps 3;

QY 13 KQQTNRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPAFRDSSENEIAV 71
DB 5 KKQTEDISSVYEIRKLGSGAFSEVFLVLAQERGAHLVAKCIPKALRGKALVENEIAV 64
QY 72 LKKIKHENIVTLEDIYESTHYLYVMQVSGGELEFDRILRGVYTEKDAVLVIOQVLSAV 131
DB 65 LRRISHPNVALEDVHESPSHLYLAMELVGTGELFDRIMERGSYTEKDAHLVGQVLGAV 124
QY 132 KYLHENGIVHRDLKPENLYLTPENSKIMITDFGLSKMEQNGIMSTACGTPGYVAPEVL 191
DB 125 SYLHSLGIVHRDLKPENLYLATPFEDSKIMVSDFGLSKIQAGNMLGTACGTPGYVAPELL 184
QY 192 AOKPYSKAVDCWSIGVITVILLCGYPFPFYETESKLFKEKIGEVFEFSPFDDISESAK 251
DB 185 EQKPYKAVDVWALGVISYILLCGYPFPFYDESDPELFSQILRASVEFSPFDDISESAK 244
QY 252 DFICHLLEKDPNERYTCEKALSHPWIDGNLTALHRDIYFVSLSIQIKNFASKWROAFNAA 311
DB 245 DFIRHLLEDPKQRTCCQALQHLWISGDAFDRDILGSVSIQIKNFARHAKRAFNAF 304
QY 312 AVVHMRKLMHNLHSPGVPRPEVENRPPETQASSETSRPSPPEITITEAP 359
DB 305 SFLRHIRKL-----GQSPGE-----EASRQCMTRHSHPLGLTQSP 341

RESULT 11
Q9NNY2 PRELIMINARY; PRT; 343 AA.
ID Q9NNY2
```

```
AC Q9NNY2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CA2+/calmodulin-dependent protein kinase I.
GN CAMK1.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97432815; PubMed=9286695;
RA Brenner V., Nyakatura G., Rosenthal A., Platzer M.;
RT "Genomic organization of two novel genes on human Xq28: compact head
RT to head arrangement of IDH gamma and TRAP delta is conserved in rat
RT and mouse.";
RL Genomics 44:8-14(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Platzer M.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; U52111; AAF74509.1; -.
DR HSSP; Q63450; 1A06.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 343 AA; 38500 MW; E954AE89DE508C9 CRC64;

Query Match 41.7%; Score 1047; DB 4; Length 343;
Best Local Similarity 60.7%; Pred. No. 1.2e-74;
Matches 201; Conservative 52; Mismatches 70; Indels 8; Gaps 2;

QY 13 KQQTNRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPAFRDSSENEIAV 71
DB 5 KKQTEDISSVYEIRKLGSGAFSEVFLVLAQERGAHLVAKCIPKALRGKALVENEIAV 64
QY 72 LKKIKHENIVTLEDIYESTHYLYVMQVSGGELEFDRILRGVYTEKDAVLVIOQVLSAV 131
DB 65 LRRISHPNVALEDVHESPSHLYLAMELVGTGELFDRIMERGSYTEKDAHLVGQVLGAV 124
QY 132 KYLHENGIVHRDLKPENLYLTPENSKIMITDFGLSKMEQNGIMSTACGTPGYVAPEVL 191
DB 125 SYLHSLGIVHRDLKPENLYLATPFEDSKIMVSDFGLSKIQAGNMLGTACGTPGYVAPELL 184
QY 192 AOKPYSKAVDCWSIGVITVILLCGYPFPFYETESKLFKEKIGEVFEFSPFDDISESAK 251
DB 185 EQKPYKAVDVWALGVISYILLCGYPFPFYDESDPELFSQILRASVEFSPFDDISESAK 244
QY 252 DFICHLLEKDPNERYTCEKALSHPWIDGNLTALHRDIYFVSLSIQIKNFASKWROAFNAA 311
DB 245 DFIRHLLEDPKQRTCCQALQHLWISGDAFDRDILGSVSIQIKNFARHAKRAFNAF 304
QY 312 AVVHMRKLMHNLHSPGVPRPEVENRPPETQASSETSRPSPPEITITEAP 342
DB 305 SFLRHIRKL-----GQIPGECAEQGMA 328

RESULT 12
Q9TXJ0 PRELIMINARY; PRT; 348 AA.
ID Q9TXJ0
AC Q9TXJ0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
```

```
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical 39.1 kDa protein.
GN K07A9.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Davidson S., O'Neal D.;
RT "The sequence of C. elegans cosmid K07A9.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF099924; AAF23187.1; -.
DR HSSP; Q63450; 1A06.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Hypothetical protein; Serine/threonine-protein kinase;
KW Transferase.
SQ SEQUENCE 348 AA; 39124 MW; 88525C39080A709F CRC64;

Query Match 40.1%; Score 1008; DB 5; Length 348;
Best Local Similarity 57.3%; Pred. No. 1.5e-71;
Matches 185; Conservative 68; Mismatches 68; Indels 2; Gaps 2;

QY 19 IRKTFIPEVLGSGAFSEFLVKQRL-TGKLFALKCI-KKSPAFRDSLSLENIATVLKIK 76
Db 18 IREKYDFRDLGTGAFSKVFLAESKSDAGQMYAVKCIDKALKGKEESLENIKVLRL 77

QY 77 HENITVLEDIYESTHYLYMQLVSGELFDRILRGVYTEKDSALVIOQVLSAVKYLHE 136
Db 18 IREKYDFRDLGTGAFSKVFLAESKSDAGQMYAVKCIDKALKGKEESLENIKVLRL 77

QY 137 NGIVHRDLKPENLLYLTPEENSKIMTDFGLSKMEQNGIMSTACGTPGYVAPVLAQPY 196
Db 18 HNNIVQLFDIYDEKQFYVLMELVTGGELFDRIVAKSGSYTEQDASNLIRQVLEAVGFMD 137

QY 197 SKAVDCWSIGVITVILLCGYPPEYTESKLFKEIKEGYEFESPFWDIDISAKDFICH 256
Db 18 HNNIVQLFDIYDEKQFYVLMELVTGGELFDRIVAKSGSYTEQDASNLIRQVLEAVGFMD 137

QY 198 GKAVDWSIGVIAYILLCGYPPEYTESKLFKEIKEGYEFESPFWDIDISAKDFICH 257
Db 18 HNNIVQLFDIYDEKQFYVLMELVTGGELFDRIVAKSGSYTEQDASNLIRQVLEAVGFMD 137

QY 137 NGIVHRDLKPENLLYLTPEENSKIMTDFGLSKMEQNGIMSTACGTPGYVAPVLAQPY 196
Db 18 HNNIVQLFDIYDEKQFYVLMELVTGGELFDRIVAKSGSYTEQDASNLIRQVLEAVGFMD 137

QY 197 SKAVDCWSIGVITVILLCGYPPEYTESKLFKEIKEGYEFESPFWDIDISAKDFICH 256
Db 18 HNNIVQLFDIYDEKQFYVLMELVTGGELFDRIVAKSGSYTEQDASNLIRQVLEAVGFMD 137

QY 198 GKAVDWSIGVIAYILLCGYPPEYTESKLFKEIKEGYEFESPFWDIDISAKDFICH 257
Db 18 HNNIVQLFDIYDEKQFYVLMELVTGGELFDRIVAKSGSYTEQDASNLIRQVLEAVGFMD 137

QY 257 LLEKDPNERYTCEKALSHPWIDGNLTALHRDIYPSVLSQIOKNFAKSKWRQAFNAAVVHH 316
Db 18 HNNIVQLFDIYDEKQFYVLMELVTGGELFDRIVAKSGSYTEQDASNLIRQVLEAVGFMD 137

QY 258 LMCCDPEARFTCDALSHPWISGNTAYTHDHTGTVAVHLKSLAKRNNKKAYNAAAIAIQ 317
Db 18 HNNIVQLFDIYDEKQFYVLMELVTGGELFDRIVAKSGSYTEQDASNLIRQVLEAVGFMD 137

QY 317 MRKLHMLHSPGVPRPEVNRPE 339
Db 18 HNNIVQLFDIYDEKQFYVLMELVTGGELFDRIVAKSGSYTEQDASNLIRQVLEAVGFMD 137

QY 318 LQMLRLSSNRLQKQASQOQPE 340
Db 18 HNNIVQLFDIYDEKQFYVLMELVTGGELFDRIVAKSGSYTEQDASNLIRQVLEAVGFMD 137

RESULT 13
Q9UAH6
```

```
ID Q9UAH6 PRELIMINARY; PRT; 348 AA.
AC Q9UAH6;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Ca2+/calmodulin-dependent protein kinase I.
GN CMK-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99357789; PubMed=10428833;
RA Eto K., Takahashi N., Kimura Y., Masuho Y., Arai K., Muramatsu M.,
RA Tokumitsu H.;
RT "Ca2+/calmodulin-dependent protein kinase cascade in Caenorhabditis
RT elegans. Implication in transcriptional activation.";
RL J. Biol. Chem. 274:22556-22562(1999).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB021864; BAA82674.1; -.
DR HSSP; Q63450; 1A06.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 348 AA; 39066 MW; 24A2633BF4D88DB0 CRC64;

Query Match 40.0%; Score 1004; DB 5; Length 348;
Best Local Similarity 57.3%; Pred. No. 3e-71;
Matches 185; Conservative 67; Mismatches 69; Indels 2; Gaps 2;

QY 19 IRKTFIPEVLGSGAFSEFLVKQRL-TGKLFALKCI-KKSPAFRDSLSLENIATVLKIK 76
Db 18 IREKYDFRDLGTGAFSKVFLAESKSDAGQMYAVKCIDKALKGKEESLENIKVLRL 77

QY 77 HENITVLEDIYESTHYLYMQLVSGELFDRILRGVYTEKDSALVIOQVLSAVKYLHE 136
Db 18 HNNIVQLFDIYDEKQFYVLMELVTGGELFDRIVAKSGSYTEQDASNLIRQVLEAVGFMD 137

QY 137 NGIVHRDLKPENLLYLTPEENSKIMTDFGLSKMEQNGIMSTACGTPGYVAPVLAQPY 196
Db 18 NGVYHRDLKPENLLYLTPEENSKIMTDFGLSKMEQNGIMSTACGTPGYVAPVLAQPY 197

QY 197 SKAVDCWSIGVITVILLCGYPPEYTESKLFKEIKEGYEFESPFWDIDISAKDFICH 256
Db 18 GKAVDWSIGVIAYILLCGYPPEYTESKLFKEIKEGYEFESPFWDIDISAKDFICH 257

QY 198 GKAVDWSIGVIAYILLCGYPPEYTESKLFKEIKEGYEFESPFWDIDISAKDFICH 257
Db 18 GKAVDWSIGVIAYILLCGYPPEYTESKLFKEIKEGYEFESPFWDIDISAKDFICH 257

QY 257 LLEKDPNERYTCEKALSHPWIDGNLTALHRDIYPSVLSQIOKNFAKSKWRQAFNAAVVHH 316
Db 18 LMCCDPEARFTCDALSHPWISGNTAYTHDHTGTVAVHLKSLAKRNNKKAYNAAAIAIQ 317

QY 317 MRKLHMLHSPGVPRPEVNRPE 339
Db 18 LQMLRLSSNRLQKQASQOQPE 340

RESULT 14
Q9V314 PRELIMINARY; PRT; 405 AA.
ID Q9V314 Q9V4D3;
AC Q9V314; Q9V4D3;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE CamKI protein.
GN CAMKI OR CG1495.
OS Drosophila melanogaster (Fruit fly).
```

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=RETINA;  
RC MEDLINE=99030403; PubMed=9813038;  
RA Xu X.Z.S., Wes P.D., Chen H., Li H.S., Yu M., Morgan S., Liu Y.,  
RA Montell C.,  
RT "Retinal targets for calmodulin include proteins implicated in  
RT synaptic transmission.";  
RL J. Biol. Chem. 273:31297-31297(1998).  
RN [2]  
RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers J., H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Flossler A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-f., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL: Y17917; CAA76937.1; -;  
DR EMBL: AE003844; AAF59343.1; -;  
DR EMBL: AE003844; AAF59344.1; -;  
DR HSSP: Q63450; IA06.  
DR FlyBase: FBgn0016126; CamKI.  
DR InterPro: IPR000719; Ser\_pkinase.  
DR InterPro: IPR002290; Ser\_thr\_pkinase.  
DR Pfam: PF00069; pkinase; 1.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00220; S\_TKc; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS0108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Alternative splicing; Kinase;  
KW Serine/threonine-protein kinase; transferase.

FT VARSPLIC 171 190 PENLLYSPDDDDSKIMISDF -> VRQFYLIYFNHDLALR  
FT FT KGK (IN SHORT ISOFORM).  
FT VARSPLIC 191 405 MISSING (IN SHORT ISOFORM).  
SQ SEQUENCE 405 AA; 45113 MW; BDE60C6C981CFBC4 CRC64;  
Query Match 38.1%; Score 957.5; DB 5; Length 405;  
Best Local Similarity 55.4%; Pred. No. 1.8e-67;  
Matches 190; Conservative 64; Mismatches 70; Indels 19; Gaps 5;  
QY 2 GRK-EEDDCSSWKKOTTNIRKTFEIMFVLGSGAFSEFLVKQR-LTGKLFALKCI-KKSP 58  
Db 10 GKAKAKDLKELNQ-VSIEKYNHGLLGTGAFSEVRLAESKDPGEHFAVKIIDKAL 68  
QY 59 AFRDSSLENEIAVLK-----KIKHENITVLTEDIYESTHYLLVMQLVSGG 103  
Db 69 KGKEESLENEIRVLRFSANHFDCGLNGTRLTHPNVQLLETYEDSKSVLVNMLVTGG 128  
QY 104 ELFDRIILRGVYTKDASLVQVLSAVKYLHENGIVHRLDKPENLILYLPENSKIMIT 163  
Db 129 ELFDRIIVEKSGYTEKDASHLIRLEADVNMHEQGVVHRLDKPENLILYSPDDSKIMIS 188  
QY 164 DFLSKMEONGIMTAGTTCGYVAPEVLAQPKYKAVDCWSIGVITVILLCGYPPFYEE 223  
Db 189 DFLSKMEDSGIMATACGTCGYVAPEVLAQPKYKAVDCWSIGVITVILLCGYPPFYDEN 248  
QY 224 ESKLFEKIKEGYEFESPFWDDISAKDFICHLLERKDPNERYTCEKALSHPWIDGNAL 283  
Db 249 DANLFAQLKGLDFEEDSPYWDSEISAKHFKNLKMCVTVEKRYTCKQALGHAMISGNEAS 308  
QY 284 HRDIYPSVLSIQNFKAKSKWRQAFNAANVHHRKLMNLHS 326  
Db 309 SRNIHGTVSEQLKKNFAKSRMKQAYATVIRQMORNALNSNS 351  
RESULT 15  
Q8WTT8 PRELIMINARY; PRT; 501 AA.  
AC Q8WTT8  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Hypothetical 54.4 kDa protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRNIN;  
RA Strausberg R.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LUNG;  
RA Strausberg R.;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC019256; AAH19256.1; -;  
DR EMBL: BC017363; AAH17363.1; -;  
DR InterPro: IPR000104; Antifreeze\_1.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002290; Ser\_thr\_pkinase.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR Pfam: PF00069; pkinase; 1.  
DR PRINTS: PR00308; ANTIFREEZE1.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00220; S\_TKc; 1.  
DR SMART: SM00219; TyrKc; 1.  
DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 501 AA; 54353 MW; ID6DBF8ABA0DD1D70 CRC64;  
Query Match 29.7%; Score 747.5; DB 4; Length 501;  
Best Local Similarity 35.9%; Pred. No. 9.4e-51;

		Matches 175; Conservative 81; Mismatches 176; Indels 55; Gaps 12;			
Qy	15	QTTNIRKTFIFMEV	LGSAFSEVFLVKQRLTGKLFALKCIKKSPAFRD-----SSLENEI	69	
Db	16	QPSVTDYDLGQV	IKTEEFCEIFRAKDKTTGKLHTCKKFKQK-----RDGRKVRKAAKNEI	71	
Qy	70	AVLKKIKHENIV	LEDIYESTHYLYVMQLVSGGELFDRILERGYVTEKDALIVTQOVL	129	
Db	72	GILKMKHPN	LQLDVDFVTRKEYFIFLELATGREVFDWILDQGYSERDTSNVVRQVLE	131	
Qy	130	AVKYLHENGIV	HRDLKPNLLYLTPEENSKIMITDFGLSKMEQNGIMSTACGTPGYVAPE	189	
Db	132	AVAYLHSLKIV	HRNKLNLVYNNRLKNSKIVISDFHLAKLE-NGLIKEPCGTPEYLAPE	190	
Qy	190	VLAQPKYKAVD	CWSIGVITYIILCGYPPPYEETE-----SKLFEKIKEGYEYFESP	241	
Db	191	VVGRQRYGRPV	DCWAIGVIMYIILSGNPPPYEEVEEDDYENHDKNLFKILAGDYEFDSP	250	
Qy	242	FWDDISESAK	DFICHLEKDPNERYTCEKALSHPWIDGNTALHROIYPSVSLQIOKNFAK	301	
Db	251	YWDDISOAAK	DVTRLMEVEQDORITAEAAISHENISGNAADKNIKOGVCAQIEKNFAR	310	
Qy	302	SKWQAFNAAV	VHHMKLHMLHSPGVRPEVENRPPETQASSTRPSPSEITITEAPVL	361	
Db	311	AKKKKAVRV	TTL---MKRL-----RAPE--QSSTAAAQASASATDTATPGA	350	
Qy	362	---DHSV	ALPALTQLPCQHGRP-----TAPGGRSLNCLVNGSL--HSSSLVPMHQGSL	411	
Db	351	AGGATAAA	AGSATSAPGDAARAAKSDNVAPADRSATPATDGSATPATDGSVTPTATDGS	410	
Qy	412	AAGPCGCC	SSCLN----IGSKGKSSYCSEPTLLKANKQNFKEVMV---PVKASGSSH	464	
Db	411	TPATDGS	VTPTATDGSATPATDGRATPATEESTVPTTQSSAMLATKAAATPEFAMAQPDST	470	
Qy	465	CRAQGTG	471		
Db	471	APEGATG	477		

Search completed: March 14, 2003, 14:09:39  
Job time : 35 secs



Db 1487 GGTGGGACACTGGGGTGTCTCTCTATGTATGATCCAGAGGCCATGTGGTTTCAGGA 1546  
QY 1528 CAATTTTCAGGACATATTTCAACTCTCTCTCTCTCCAAACCTGGTCTATATCCGGCAG 1587  
Db 1547 GACATTTCCCTG--GCTCTTCTCTGCTCTCTCTGAGCTGGCATCTGCGCTGAGGAGGGGGC 1604  
QY 1588 AGGGAGGAGGAGAGAGAGTGGAGGAGGCTTAGCAGGAGCAGTTTCTGCCAGAGCA 1647  
Db 1605 AGTGAATAGATAGCAAAAGCAGTGGTCTGAGGAG----- 1643  
QY 1648 CCAGCCTGTCTGACGGGGAGCCCTCTATAGAGGAGCCAGGAGGAGGCCAAAGCGT 1707  
Db 1644 TAGGGCTGTGTGCTGCTGGCAGACATCACAGAGAGTCCAGAGGAGCCAGCCCATCT 1703  
QY 1708 AGNAGCCTGTGTGAAGCTGTGAGCAGGAGAGAGCGGT-----GCCACACAGCT 1754  
Db 1704 TGGACGTTCTGTGAGGAGCCACAGAGGATGATGAGTGGCTGCTGCTCCAGCATCCAGCC 1763  
QY 1755 TCAGGCTCTCCCTGACCTGCTCTCTATGCCCCACACCCCTGAGGCTGGCTCTGTC 1814  
Db 1764 TCCAGTCTCCC-GACCTGCCATTCCTTCCACACATCATACGTCAGTGGCTCCATGC 1822  
QY 1815 AGTGTACGTAGTAGCTCTGCTGCTGGTCTGTGCTGTTTGTGCGTAAAGCTTAATGGC 1874  
Db 1823 AGTGTCCATAGATAGTGTCTGCTGGTCTGGTGTGTTGTTGTGTAAGGCTTTATGGC 1882  
QY 1875 TGGCCAGGCTGTGACCTCTCTCCAGCAAGGAGGATGAGGAGTACCCAGACTCCCA 1934  
Db 1883 TGGCCAGGCTGTGACCTCTCTCAAGCAAGGAGGATGAGGAGTAACTAGACTCCCTG 1942  
QY 1935 CTCTGCACACTCTACTCCCTCTCAAGCTCTCAAGCTCTTGGCCAGATGAGGCTCAT 1994  
Db 1943 CTCTGCAGACTGCT 1993  
QY 1995 TAATGTCTGTGCTGCTGCT 2054  
Db 1994 TATCATAG-TACCTGCT 2044  
QY 2055 TGAGCTCTTCAAGTTCTTAATCTTAATCTTAATCTTAATCTTAATCTTAATCTTAATCT 2114  
Db 2045 GGAGCACTCACAACCTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2104  
QY 2115 GCAGCAGACTCTCTGCT 2173  
Db 2105 GGCACATC-----CAAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2158  
QY 2174 AAGTGAATAGAAAGATTCATGATGAGGCTGCAAGGATTCATCTCTCTCTCTCTCTCT 2233  
Db 2159 TGGTGGACTAGAGAAAGTGTGACTATGAGCTGTGAGCTTCTCTCTCTCTCTCTCTCT 2218  
QY 2234 GTCTCGGTGCACACACCAAGGAGGATTAACCTTGAAGTGTGACTATTTTAACTGTGCT 2293  
Db 2219 ACTCATATATCCCTTAATTAAG-----ATTGCGGAGCTATTTCCATCTCATCTC 2267  
QY 2294 AGGAGTTCTAATCT 2353  
Db 2268 AGG-----TCT 2321  
QY 2354 TGTCTTCT 2412  
Db 2322 ----GTCCT 2376  
QY 2413 AATAAATGACATATTTTCT 2447  
Db 2377 AATAAATGACATATTTTCT 2411

RESULT 8  
LOCUS HS272L16 157875 bp DNA linear PRI 23-NOV-1999  
DEFINITION Human DNA sequence from clone 272L16 on chromosome 1q32.1-32.3.  
Contains the 3' end of the LAMB3 gene for Laminin, Beta 3 (Nicein,  
Kalinin, BM600) and a novel Rat Ca2+/Calmodulin dependent Protein

Kinase LIKE gene. Contains ESTs, STSs, GSSs, genomic marker D1S491  
and a ca repeat polymorphism, complete sequence.  
AL023754.1 GI:4007152  
HTG: BM600; ca repeat polymorphism; Ca2+/Calmodulin dependent  
Protein Kinase; D1S491; Kalinin; LAMB3; Laminin Beta 3; Nicein.  
Homo sapiens.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 157875)  
Grafham,D.  
Direct Submission  
Submitted (27-NOV-1998) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
requests: clonerequest@sanger.ac.uk  
On Dec 12, 1998 this sequence version replaced gi:3873472.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence is the entire insert of clone 272L16. This sequence  
has been finished according to sequence map criteria as follows. An  
attempt is made to resolve all sequencing problems, such as  
compressions and repeats, but not necessarily within known  
annotated human repeat sequence elements (e.g. Alu). Where the  
sequence is ambiguous, there is an annotation using the 'unsure'  
feature key.  
272L16 is from the library RPII constructed at the Roswell Park  
Cancer Institute by the group of Pieter de Jong. For further  
details see <http://bacpac.med.buffalo.edu/VECTOR:pcYPAC2> This  
sequence was generated from a human chromosome 1 bacterial clone  
contig constructed in collaboration by the Sanger Centre chromosome  
1 mapping group and Brian Schutte, Bryan Bjork, Kevin Coppage and  
Jeffrey Murray. Department of Pediatrics, University of Iowa, USA.  
Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr1>.

FEATURES  
Location/Qualifiers  
1..157875  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="q32.1-32.3"  
/clone="RP1-272L16"  
/clone\_lib="RPII-1"  
966..1078  
/note="MLTII repeat: matches 281..410 of consensus"  
1082..1506  
/note="MSTD repeat: matches 1..426 of consensus"  
1585..1875  
/note="MLTII repeat: matches 1..319 of consensus"  
1898..2024  
/note="MLTII repeat: matches 58..203 of consensus"  
2347..2925  
/note="MER34 repeat: matches 6..543 of consensus"  
3090..3417  
/note="MLTII2 repeat: matches 23..374 of consensus"  
3677..4074  
/note="MER54B repeat: matches 486..902 of consensus"  
4089..5338  
/note="MLTII2-internal repeat: matches 358..1643 of  
consensus"  
5339..5532  
/note="AluJo repeat: matches 101..302 of consensus"  
5536..5966  
/note="MLTII-internal repeat: matches 5..450 of  
consensus"  
5972..6019  
/note="MLTII repeat: matches 318..365 of consensus"  
6020..6459  
/note="LTR7 repeat: matches 1..450 of consensus"







	source	1..1013 /organism="Rattus norvegicus" /db_xref="taxon:10116" /clone="N5" /tissue_type="brain" /clone_lib="S. Nakanishi" /dev_stage="embryo (E18)" <!. .86 87.. >1013 /notes="similar to calcium/calmodulin-dependent protein kinase I mRNA (L24907, L26288)"; /codon_start=1 /product="Protein Kinase" /protein_id="BAAL9880.1" /db_xref="GI:2077934" /translation="MGRREEDCSWKQTNIRKTFPMELVSGAFSEFLVKORY TGKLFALKCIKKSPAPROSSLENEIAVLKRHNIVTLEDIYESITHYLYVMOLVS GELFDRLERGVYTBDASLVQQVSLSKYLHENGIVHRDLKPENLLYLTPEBNSKI MITDFGLSKMEQNVTMGSTGTGYVAPEVLAOKPYSKAVDCWMSIGVITVILCGPYP FYETESKLFEIKEGYYEFESPFWDDISEAKDIFCHLLEKDPNERITYTEKAHRHPW IDNTAHLRDIVPSYSLOIQKNFAKSQRQAFN"
	5'UTR	
	CDS	BASE COUNT    276 a   245 c   262 g   230 t ORIGIN
		Query Match                  33.2%; Score 813; DB 10; Length 1013; Best Local Similarity      90.2%; Pred. No. l.le-213; Matches 870; Conservative   0; Mismatches 95; Indels     0; Gaps     0;
D b 147711	CAGCTTCAGATCTCCCTGACCTGCCTGCTCTCATGGCCCCCACACCCTACGTGCCGTGGCTC	147770
O Y 1810	TGTGCAGTTACGTAGTAGACTCGCGCTGGGTCTGTGCTGTGTTGCTGAAGAAGCTTA	1869
D b 147771	TGTGCAGTTACGTAGTAGACTCGCGCTGGGTCTGTGCTGTGTTGCTGAAGAAGCTTA	147830
O Y 1870	TGGCTGGCCAGGCTGTGCTCACTCTCCAAAGCAAAGCCATATGAGCATATACCCAGAC	1929
D b 147831	TGGCTGGCCAGGCTGTGCTCACTCTCCAAAGCAAAGCCATATGAGCATATACCCAGAC	147890
O Y 1930	TCCCAGCTGCACACACTCACTCCCACTCTCAAGCCTCCAACCTCTTGGCCAGATTGGG	1989
D b 147891	TCCCAGCTGCACACACTCACTCCCACTCTCAAGCCTCCAACCTCTTGGCCAGATTGGG	147950
O Y 1990	CTCATTAATGTCTGTGCTGCCATCTGCGATGATGACAGGCGAGCTCCCATGTGGTCT	2049
D b 147951	CTCATTAATGTCTGTGCTGCCATCTGCGATGATGACAGGCGAGCTCCCATGTGGTCT	148010
O Y 2050	GCTGTGAGCTCTCAAGTTCTTAATCCTTAACTPCAGGATTAGCTGCCAAGTGGCTGAG	2109
D b 148011	GCTGTGAGCTCTCAAGTTCTTAATCCTTAACTPCAGGATTAGCTGCCAAGTGGCTGAG	148070
O Y 2110	ACCCAGCAGCACACTCTGCGCCTCTCCCTGCTCAATCTAAAAGCAGTGCACACCC	2169
D b 148071	ACCCAGCAGCACACTCTGCGCCTCTCCCTGCTCAATCTAAAAGCAGTGCACACCC	148130
O Y 2170	TCCAAAGTGGAAATAAGAAAGTTCTATGATGAGGCTGCAAGGATTTCTATCCTGGCC	2229
D b 148131	TCCAAAGTGGAAATAAGAAAGTTCTATGATGAGGCTGCAAGGATTTCTATCCTGGCC	148190
O Y 2230	ACATGCTCCTCGTGACACACCAACAATGGAGTTAACCTTGGAAAGTTGACTATTAAATGTC	2289
D b 148191	ACATGCTCCTCGTGACACACCAACAATGGAGTTAACCTTGGAAAGTTGACTATTAAATGTC	148250
O Y 2290	TGCCAGGAGTTCTAATCCTGGCTGTGTCTCCTTCTCCTTGAAGTGCAGACACCA	2349
D b 148251	TGCCAGGAGTTCTAATCCTGGCTGTGTCTCCTTCTCCTTGAAGTGCAGACACCA	148310
O Y 2350	TTCTTGCTCTCCCGAGTTTCCTCGCCCTCCACCCCTCCAGCTTCATGCTCAGTGTGTG	2409
D b 148311	TTCTTGCTCTCCCGAGTTTCCTCGCCCTCCACCCCTCCAGCTTCATGCTCAGTGTGTG	148370
O Y 2410	CTTAATAAANTGGACATATTTTCTCTA 2437	
D b 148371	CTTAATAAANTGGACATATTTTCTCTA 148398	
R E S U L T       9		
D B 6557	L O C U S                               1013 bp       m R N A       linear       ROD 07-FEB-1999	
D E F I N I T I O N	Rattus norvegicus mRNA for Protein Kinase, partial cds.	
A C C E S S I O N	D86557	
V E R S I O N	D86557.1 GI:2077933	
K E Y W O R D S	Protein kinase.	
S O U R C E	Rattus norvegicus embryo (E18) brain cdna to mRNA, clone lib:S. Nakanishi clone:N5.	
O R G A N I S M	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. 1 (bases 1 to 1013) Xokokura,H., Terada,O., Naito,Y. and Hidaka,H. Isolation and comparison of rat cDNAs encoding Ca2+/calmodulin-dependent protein kinase I isoforms Biochim. Biophys. Acta 1338 (1), 8-12 (1997)	
R E F E R E N C E	9728532	
M E D L I N E	2 (bases 1 to 1013)	
J O U R N A L	Xokokura,H. Direct Submission Submitted (15-JUL-1996) Hisayuki Yokokura, Nagoya University School of Medicine, Department of Pharmacology; Tsushima 65, Showa-ku, Nagoya, Aichi 466, Japan (Tel:052-744-2075, Fax:052-744-2083)	
A U T H O R S	Location/Qualifiers	
T I T L E		
J O U R N A L		
P U B L I S H E R S		



LOCUS HS272L161 1738 bp mRNA linear PRI 21-APR-1999  
 DEFINITION Human gene isolated from PAC 272L16, Chromosome 1, similar to  
 ACCESSION AL049688  
 VERSION AL049688.1 GI:4678721  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Rhodes.S.  
 TITLE Direct Submission  
 JOURNAL Submitted (21-APR-1999) E-mail contact: humquery@sanger.ac.uk  
 COMMENT This sequence was generated from cDNA clones isolated using  
 sequence from the bacterial clone 272L16 (AL023754) and EST data.  
 The EST sequences listed match this sequence with an identity of at  
 least 95% between the coordinates shown.  
 Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr1/ Partial, experimentally  
 determined gene.  
 Sanger Centre name: dJ272L16.Cl.1.  
 Location/Qualifiers  
 1..1738  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /map="1q32.1-32.3"  
 /cl.1448  
 /codon\_start=3  
 /product="hypothetical protein"  
 /protein\_id="CAB41259.1"  
 /db\_xref="GI:4678722"  
 /translation="STLEAMGRKEEDDCSSWKKOTTNIRKTFIPMEVLGSGAFSEVFL  
 VKQRTGLKLFALKICKSPAFRDSLENIATLKKIKHENIVTLEDIYESTHYILVM  
 OLVSSEGLFDRILRGVYTKDASLVTPQVLSAVKYLHENGIVHRDLKPNLLLYLTP  
 ENSKIMIDFDGLSKMEQNGINSTAGTPGYVAPVLAQKPYSKAVDCWSIGVTYILL  
 GYPPEETESKLFPEIKEGYVEFSPFWDIDISEAKDFICHLLEKDPNERVCEKA  
 LSHPWIDGNTALHDIYPSVLSIOIKNPAKSKWRQAFNAAVVHHMKRLHMLHSPGV  
 RPEVNRPEPQASSETSRPSFEITITAPVLDHRSVALPALTPQLPCQHQHRRTPAPGR  
 SLNCLVNGSLHISLSPVPMHOGSLAAGPCGCCSCLNIGSKGSSYSCSEPTLLKKANK  
 KONFKSEYVMVPVKSGSSHCRAGGTGVCILIM"  
 <1..107  
 /number=1  
 exon  
 108..236  
 /number=2  
 exon  
 237..311  
 /number=3  
 exon  
 312..450  
 /number=4  
 exon  
 451..574  
 /number=5  
 exon  
 575..650  
 /number=6  
 exon  
 651..763  
 /number=7  
 exon  
 764..842  
 /number=8  
 exon  
 843..930  
 /number=9  
 exon  
 931..1355  
 /number=10  
 exon  
 1356..1483  
 /number=11  
 exon  
 1484..1736  
 /number=12  
 exon  
 432 a 491 c 458 g 357 t

## ORIGIN

Query Match 70.9%; Score 1734.8; DB 9; Length 1738;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1736; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 53 CTTCAACTCTGGAGGCAATGGGTGCGAAAGGAAGAAGATGACTGCAGTTCTTGGAGAAGAAC 112  
 DB 1 CTTCAACTCTGGAGGCAATGGGTGCGAAAGGAAGAAGATGACTGCAGTTCTTGGAGAAGAAC 60  
 QY 113 AGACCACCAACATCCGGAACACCTTCATTTTATGGAAGTGTCTGGATCAGGAGCTTTCT 172  
 DB 61 AGACCACCAACATCCGGAACACCTTCATTTTATGGAAGTGTCTGGATCAGGAGCTTTCT 120  
 QY 173 CAGAAGTTTTCCTGGTGAAGCAAAAGACTGACTGGGAAGCTCTTTGCTGCAAGTGCATCA 232  
 DB 121 CAGAAGTTTTCCTGGTGAAGCAAAAGACTGACTGGGAAGCTCTTTGCTGCAAGTGCATCA 180  
 QY 233 AGAAGTCACTGCTTCCGGGACAGCAGCTCGAGATGAGATTGCTGTGTTGAAAAAGA 292  
 DB 181 AGAAGTCACTGCTTCCGGGACAGCAGCTCGAGATGAGATTGCTGTGTTGAAAAAGA 240  
 QY 293 TCAAGCATGAAACATTTGACCTCGAGGACATCTATGAGACACACCCACTACTACC 352  
 DB 241 TCAAGCATGAAACATTTGACCTCGAGGACATCTATGAGACACACCCACTACTACC 300  
 QY 353 TGGTCATGACAGTTGTTCTGTTGGGAGCTCTTTGACCGGATCCCTGGAGCGGGGTGTCT 412  
 DB 301 TGGTCATGACAGTTGTTCTGTTGGGAGCTCTTTGACCGGATCCCTGGAGCGGGGTGTCT 360  
 QY 413 ACACAGAGAAGGATGCCAGCTCTGGTATCCAGCAGCTCTTGTGCGGAGTGAATACCTAC 472  
 DB 361 ACACAGAGAAGGATGCCAGCTCTGGTATCCAGCAGCTCTTGTGCGGAGTGAATACCTAC 420  
 QY 473 ATGAGATGCGATCGTCCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTG 532  
 DB 421 ATGAGATGCGATCGTCCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTG 480  
 QY 533 AAGAACTCTAAGATCATGATCAGTCTTGGTGTCTGTCACAGATGGACAGAAATGGCA 592  
 DB 481 AAGAACTCTAAGATCATGATCAGTCTTGGTGTCTGTCACAGATGGACAGAAATGGCA 540  
 QY 593 TCATGTCACCTGCTGTGGGACCCCGAGGCTACGTGGTCCAGAAAGTGTGGCCCGAAGAAC 652  
 DB 541 TCATGTCACCTGCTGTGGGACCCCGAGGCTACGTGGTCCAGAAAGTGTGGCCCGAAGAAC 600  
 QY 653 CCTACAGAAAGGCTGTGGATTGCTGCTCCATCGGCGTCATCACTACATATTGCTCTGTG 712  
 DB 601 CCTACAGAAAGGCTGTGGATTGCTGCTCCATCGGCGTCATCACTACATATTGCTCTGTG 660  
 QY 713 GATACCCCGCTTCTATGAAGAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCT 772  
 DB 661 GATACCCCGCTTCTATGAAGAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCT 720  
 QY 773 ACTATGAGTTTGTAGTCTCCATTTCTGGGATGACATTTTCTGAGTCAGCAAGGACTTTATT 832  
 DB 721 ACTATGAGTTTGTAGTCTCCATTTCTGGGATGACATTTTCTGAGTCAGCAAGGACTTTATT 780  
 QY 833 GCACACTTGTGAGAGGATCCGAAAGCGGTCACCTGTGTGAGAGGCTTGTGAGTCATC 892  
 DB 781 GCACACTTGTGAGAGGATCCGAAAGCGGTCACCTGTGTGAGAGGCTTGTGAGTCATC 840  
 QY 893 CCTGGATTGACGGAACACAGCCCTCCACCGGACATCTACCCATCAGTCAGCTCCAGA 952  
 DB 841 CCTGGATTGACGGAACACAGCCCTCCACCGGACATCTACCCATCAGTCAGCTCCAGA 900  
 QY 953 TCCAGAGAACTTTGTCTAAGCAAGTGGAGCAAGCTTCAACGACGACAGCTGTGGTGC 1012  
 DB 901 TCCAGAGAACTTTGTCTAAGCAAGTGGAGCAAGCTTCAACGACGACAGCTGTGGTGC 960  
 QY 1013 ACCACATGAGAGCTACATCATGAACCTGCACAGCCGGCGCTCCGCCACAGAGTGAGA 1072  
 DB 961 ACCACATGAGAGCTACATCATGAACCTGCACAGCCGGCGCTCCGCCACAGAGTGAGA 1020



QY 1073 ACAGCCGCCCTGAACCTCAAGCCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCAGCA 1132  
Db 1021 ACAGCCGCCCTGAACCTCAAGCCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCAGCA 1080  
QY 1133 TCACCGAGGACCTGCTCTGTGACACACAGTGTAGCAGCTCCCTGCGGACCCCAATACCT 1192  
Db 1081 TCACCGAGGACCTGCTCTGTGACACACAGTGTAGCAGCTCCCTGCGGACCCCAATACCT 1140  
QY 1193 GCCAGATGGCGCGCGGCCACTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1252  
Db 1141 GCCAGATGGCGCGCGGCCACTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200  
QY 1253 GCTCCCTCCACATCAGCAGCAGCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1312  
Db 1201 GCTCCCTCCACATCAGCAGCAGCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260  
QY 1313 CTTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1372  
Db 1261 CTTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
QY 1373 CTGAGCCACACTCTCTCAAAAGGCCCAACAAACAGAACTTCAAGTGGAGGAGGAGGAGG 1432  
Db 1321 CTGAGCCACACTCTCTCAAAAGGCCCAACAAACAGAACTTCAAGTGGAGGAGGAGGAGG 1380  
QY 1433 TACAGTTAAAGCAGTGGAGTCCACCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1492  
Db 1381 TACAGTTAAAGCAGTGGAGTCCACCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440  
QY 1493 TTATGTGATCTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1552  
Db 1441 TTATGTGATCTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
QY 1553 CCTCTGCTCTTCCAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1612  
Db 1501 CCTCTGCTCTTCCAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560  
QY 1613 CAGGCTTACGAGGAGGAGTTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1672  
Db 1561 CAGGCTTACGAGGAGGAGTTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620  
QY 1673 CCTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1732  
Db 1621 CCTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680  
QY 1733 GGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1790  
Db 1681 GGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1738

RESULT 6  
BC021840  
LOCUS BC021840 2416 bp mRNA linear ROD 07-AUG-2002  
DEFINITION Mus musculus, clone MGC:30513 IMAGE:4502479, mRNA, complete cds.  
ACCESSION BC021840  
VERSION BC021840.1 GI:18256866  
KEYWORDS MGC.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 2416)  
Direct Submission  
Strausberg, R.  
Submitted (18-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Iroom, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK plate: 41 Row: 1 Column: 13  
This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction.

FEATURES  
source

Location/Qualifiers  
1..2416  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="MGC:30513 IMAGE:4502479"  
/tissue\_type="Eye, retina, mouse strain C57Bl/6"  
/clone\_lib="NIH\_MGC\_94"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
87..1520  
/codon\_start=1  
/product="Unknown (protein for MGC:30513)"  
/protein\_id="AAH21840.1"  
/db\_xref="GI:18256867"  
/translation="MGRKEEDSSWKQNTNIRKTFIMVLGSGAFSEVFLVKORV  
TGKLPALCKIKKPAFRDLSLENIIVLKRKHENIVTLEDIYETHLYLVNQLVSG  
GELFRIERGVYTKDASLVIOQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKI  
MTTDFGLSKMEQNGVMSTACGTPGVAPVLAQPKYSKAVDCWSIGVITVILCGYPP  
FYETESKLFEKIKEGYVEFPMDIIPESAKDFICHELKEDPNERYTCEKALRHPW  
IDGNTALHRDIYPSVSLQIQKFAKSKRQAFNAAVYHMRKLHNHLASPSVQVE  
NRPVSPAFEPSPDSDSSITEAPILDPTPLPRLFCSSHSPSPSPSGGSLNC  
LVNGSLRISLSPMQGLATGFCPCSSCLNIGNKRSYCSYSEPTLFRANKKONF  
KSENVPVKAGGTHCRGGGTGCLVM"

CDS

BASE COUNT 572 a 694 c 602 g 548 t  
ORIGIN

Query Match 54.3%; Score 1328.6; DB 10; Length 2416;  
Best Local Similarity 76.1%; Pred No. 0;  
Matches 1853; Conservative 0; Mismatches 494; Indels 88; Gaps 14;

QY 31 GAGTCCCTGGCATCTCAGAGAGTTCACCTCTGGAGGCAATGGTTCGAAAGAAAGAT 90  
Db 48 GAGCCCTGGCTCTCAGGAGGCTTCACTCTGGAGGAGGAGGAGGAGGAGGAGGAG 107  
QY 91 GACTGCAGTCTCTGGAG 150  
Db 108 GACTGCAGTCTCTGGAG 167  
QY 151 GTGCTGGGATCAGAGAGCTTCTCAGAGAGTTCCTGGTGAAGCAAGAGAGAGAGAG 210  
Db 168 GTGCTGGGATCAGAGAGCTTCTCAGAGAGTTCCTGGTGAAGCAAGAGAGAGAGAG 227  
QY 211 CTCTTTGCTCTGAAGTGCATCAGAGAGTCCCTGCTCCGGAGAGAGAGAGAGAGAT 270  
Db 228 CTCTTTGCTCTGAAGTGCATCAGAGAGTCCCTGCTCCGGAGAGAGAGAGAGAGAT 287  
QY 271 GAGATTGCTGTGTTGAAAGAGTCAAGCATGAAAGAGTTCACCTCGAGAGAGATCTAT 330  
Db 288 GAGATGCTGTGTTGAAAGAGTCAAGCATGAGAGATTCGACCTCGAGAGAGATCTAT 347  
QY 331 GAGAGACACACACCTACCTGCTGATGAGAGTTCCTGCTGGGAGAGAGAGAGAGAG 390  
Db 348 GAGAGACACACACCTACCTGCTGATGAGAGTTCCTGCTGGGAGAGAGAGAGAGAG 407  
QY 391 CGGATCTGAG 450  
Db 408 CGGATCTGAG 467







QY 301 LysSerLysTTPATGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu 320  
 |||||  
 Db 918 AGAGCAAGTGGAGGAGGCTTCAACGAGCAGCTGGTGGTCCACCATGAGGAAGCTA 977  
 |||||  
 QY 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProGluThr 340  
 |||||  
 Db 978 CACATGAACCTGCAACGCCGCGCTCCGCCAGAGGTGGAGAACAGCGCGCTGAACT 1037  
 |||||  
 QY 341 GlnAlaSerGluThrSerArgProSerProGluThrLeuThrLeuThrGluAlaProVal 360  
 |||||  
 Db 1038 CAAGCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCACCATCAGGAGGACCTGTC 1097  
 |||||  
 QY 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380  
 |||||  
 Db 1098 CTGGACCAACAGTGTAGCAGCTCCCTGCCCTGAGCCCAATACCTGCCAGCATGCCGCCCG 1157  
 |||||  
 QY 381 ProThrAlaProGlyArgSerLeuAsnGlySerLeuValAsnGlySerLeuHisIleSer 400  
 |||||  
 Db 1158 CCACCTGCCCTGGTGGCAGGTCCCTCAACTGCTGGTCAATGCTCCCTCCACATCAGC 1217  
 |||||  
 QY 401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420  
 |||||  
 Db 1218 AGCAGCTGGTGGCCATGATCAGAGGTCCCTGGCGCGCGGCTGTGCTGCTGCTC 1277  
 |||||  
 QY 421 SerCysLeuAsnIleGlySerLysGlyLysSerSerTyrcysSerGluProThrLeuLeu 440  
 |||||  
 Db 1278 AGCTCCCTGAAATGCGGAGCAAGAAAGTCTCTACTGCTGAGCCCAACATCCTC 1337  
 |||||  
 QY 441 LysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAlaSer 460  
 |||||  
 Db 1338 AAAAGGCGCAACAAAAACAGAACTTCAAGTCGGAGGTGATGTTACAGTTAAGCCAGT 1397  
 |||||  
 QY 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476  
 |||||  
 Db 1398 GGCAGCTCCACTGCCGGGAGGAGAGTGGAGTGTGCTGCTCATATG 1445  
 |||||

RESULT 2  
 AX399682  
 LOCUS AX399682 2447 bp DNA linear PAT 06-JUN-2002  
 DEFINITION Sequence 3 from Patent WO0224947.  
 ACCESSION AX399682  
 VERSION AX399682.1 GI:21335455  
 KEYWORDS  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1

REFERENCE  
 AUTHORS Delaney, A.D. and Yoganathan, T.  
 TITLE Cancer associated protein kinases and their uses  
 JOURNAL Patent: WO 0224947-A 3 28-MAR-2002;  
 KINETEK PHARMACEUTICALS INC (CA); UNIV BRITISH COLUMBIA (CA)

FEATURES  
 Location/Qualifiers  
 1..2447

source  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 70..1500  
 /note="unnamed protein product"

CDS  
 /codon\_start=1  
 /protein\_id="CAD33521.1"  
 /db\_xref="GI:21335456"

/translation="MGRKEEDSCSWKKOTNIRKTFIPMEVLGSAFSEFLVKRL  
 TGLFKALKIKKSPAFRDSLENIATVKKIKHENIVTLEDIYESTHYLVMLVSG  
 GELFDRLRGVYTERDASLVQVLSAVKYLHENGIVHRDLKPNLNYLTPENSKI  
 MITDFGLKMEQNGIMSTACGTPGVAPVLAOKPYKSAVDCWSIGVITYILLGYP  
 FYETESKLFKEIKGVIYEFESPDWDIISAKDFICHLEKDPNRYCTCKALSHWP  
 IDGNTALRDYIPSVSLQIKNFASKWROAFNAAVVHMRLKLNHNLSCVPRPEVE  
 NRPEQASSETSPSSPETITFAPVLDHVALPALTQLPCQGRPPAPGRSLNCL  
 VNSLHISLSPMQLAGPCCGCCSLNIGSKGKSTCYSEPTLLKANKKUNF  
 SEVHPVPAKSGSHCHRGQTGVCLIM"

BASE COUNT 590 a 707 c 604 g 546 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3,52e-170 Length: 2447  
 Score: 2513.00 Matches: 476  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Indels: 0  
 DB: 6 Gaps: 0

US-09-960-643-2 (1-476) x AX399682 (1-2447)

QY 1 MetGlyArgLysGluLysLysSerSerTrpLysGlnThrThrAsnIleArg 20  
 |||||  
 Db 70 ATGGTCTGAAAGAGAGAGATGACTGCAGTCTCTGGAAGAAACACCAACATCCGG 129  
 |||||  
 QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
 |||||  
 Db 130 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAAGTTTCTCGTG 189  
 |||||  
 QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60  
 |||||  
 Db 190 AAGCAAAAGACTGACTGGAGAGCTCTTTCCTCTGAACTGCATCAAGAACTACCTGCCTTC 249  
 |||||  
 QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80  
 |||||  
 Db 250 CGGAGACAGACCTGGAGAAATGAGATGCTGTGTTGAAAAGATCAAGCATGAACATT 309  
 |||||  
 QY 81 ValThrLeuGluAspIleTyrgluSerThrThrHisTyrrTyrrLeuValMetGlnLeuVal 100  
 |||||  
 Db 310 GTGACCTTGGAGGACATCTATGAGAGCACCACCACTACTACCTGCTCATGCAGCTTGT 369  
 |||||  
 QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrrThrGluLysAspAla 120  
 |||||  
 Db 370 TCTGTGGGAGACTCTTTGACCGATCTTGAGCGGGGTCTACACAGAAAGATGCC 429  
 |||||  
 QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrrLeuHisGluAsnGlyIleVal 140  
 |||||  
 Db 430 AGCTCGGTGATCCAGCAGGTCTTGTGCGCAGTGAATACCTACATGAGAATGGCATCGTC 489  
 |||||  
 QY 141 HisArgAspLeuLysProGluAsnLeuTyrrLeuThrProGluGluAsnSerLysIle 160  
 |||||  
 Db 490 CACAGAGACTTAAAGCCCGAAACCTGCTTACCTTACCTTACCTTACCTTACCTTACCTT 549  
 |||||  
 QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180  
 |||||  
 Db 550 ATGATCATGACTTGTGCTGCTCCAGATGGAACACAAATGGCATCATGCTCCATGCCCTGT 609  
 |||||  
 QY 181 GlyThrProGlyTyrrValAlaProGluValLeuAlaGlnLysProTyrrSerLysAlaVal 200  
 |||||  
 Db 610 GGGACCCAGGCTACGTGGCTCCAGAAAGTCTGGCCAGAAACCTTACAGCAAGGCTGTG 669  
 |||||  
 QY 201 AspCysTyrSerIleGlyValIleThrTyrrIleLeuLeuCysGlyTyrrProPheTyrr 220  
 |||||  
 Db 670 GATTGCTGCTCATCGGCGCTCATCACCTACATATTGCTCTGTGTGATATCCCGCTCTAT 729  
 |||||  
 QY 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrrTyrrGluPheGluSer 240  
 |||||  
 Db 730 GAAGAAACGGAGCTTAAGCTTTTCGAGAAGATCAGAGAGGCTACTATGATGTTGAGTCT 789  
 |||||  
 QY 241 ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260  
 |||||  
 Db 790 CCATTCGTGGATGACATTTCTGAGTACAGCAAGGACTTTTATTTGCCACTTCTGGAAG 849  
 |||||  
 QY 261 AspProAsnGluArgTyrrThrCysGluLysAlaLeuSerHisProTyrrIleAspGlyAsn 280  
 |||||  
 Db 850 GATCCGAAGCAGCGGTACACTGCTGGAGAGCCCTTCAGTCATCCCTGGATGACGGAAC 909  
 |||||  
 QY 281 ThrAlaLeuHisArgAspIleTyrrProSerValSerLeuGlnIleGlnLysAsnPheAla 300  
 |||||  
 Db 910 ACGGCCCCCAGCGGAGCATCTACCCATCAGTCCAGCTCCAGTCCAGTCCAGTCCAGTCCAGT 969  
 |||||  
 QY 301 LysSerLysTyrrArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu 320  
 |||||  
 Db 970 AAGAGCAAGTGGAGGCAAGCCTTCAACGCGCAGCAGCTGTGTGTCACCATCAGGAAAGCTA 1029  
 |||||











GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 14, 2003, 17:36:01 ; Search time 328 Seconds  
(without alignments)  
3268.144 Million cell updates/sec

Title: US-09-960-643-2

Perfect score: 2513  
Sequence: 1 MGRKEDDCSSWKQTTNIR.....VKASGSHCRAGTGVCLIM 476

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 6.0 , Ygapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DBV=xlh  
-O/cgn2\_1/USPTO.spool/US09960643/runat\_07032003\_090517\_19867/app.query.fasta\_1.647  
-DB=N\_Geneseq\_101002 -QPMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09960643@cgn1\_1\_263@runat\_07032003\_090517\_19867 -NCPU=6 -ICPU=3  
-NO\_XLIFY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -FPGAEXT=7  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FPGAEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: N\_Geneseq\_101002.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
25: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2513	100.0	1956	22	Human polynucleoti
2	2513	100.0	2447	24	Human calmodulin k
3	2499.5	99.5	2165	22	Human polynucleoti
4	2327.5	92.6	2689	22	Human diagnostic a
5	1250.5	49.8	1736	22	Human kinase (PKIN
6	1246	49.6	1074	22	Polynucleotide seq
7	1246	49.6	1074	22	Nucleotide sequenc
8	1246	49.6	1158	22	Nucleotide sequenc
9	1246	49.6	1565	24	Human Cam kinase I
10	1246	49.6	1661	22	Human kinase (PKIN
11	1246	49.6	1671	22	Nucleotide sequenc
12	1246	49.6	2164	22	Human polynucleoti
13	1243	49.5	1733	22	Human polynucleoti
14	1238.5	49.3	1694	22	Human polynucleoti
15	1237	49.2	1578	22	Human death domain
16	1150.5	45.8	1728	21	CDNA sequence enco
17	1070	42.6	1458	19	Calmodulin-depende
18	1041	41.4	1032	22	Polynucleotide seq
19	1036	41.2	1512	23	DNA encoding novel
20	990	39.4	1282	20	Disease associated
21	957.5	38.1	2264	23	Drosophila melanog
22	753	30.0	2494	22	Human full-length
23	753	30.0	2747	22	Human polynucleoti
24	753	30.0	2897	22	Human protein phos
25	753	30.0	2905	22	Human full-length
26	747.5	29.7	2297	22	Nucleotide sequenc
27	747.5	29.7	3017	22	Human kinase (PKIN
28	747.5	29.7	3124	24	Human protein kin
29	747	29.7	2840	22	Human full-length
30	733.5	29.2	3022	22	Human full-length
31	726.5	28.9	1572	22	Human protein enco
32	683.5	27.2	1774	23	DNA encoding novel
33	683.5	27.2	1774	23	DNA encoding novel
34	667	26.5	1763	23	DNA encoding novel
35	664.5	26.4	3471	18	PSKH-1 cDNA. Homo
36	661.5	26.3	1764	23	DNA encoding novel
37	650.5	25.9	1386	22	Human polynucleoti
38	650.5	25.9	1450	22	Human polynucleoti
39	639.5	25.4	3552	24	Human serine/threo
40	633.5	25.2	2667	24	Human PKIN-14 cDNA
41	625.5	24.9	2301	24	Human serine/threo
42	615.5	24.5	2110	21	CDNA sequence enco
43	611.5	24.3	1803	22	CDNA encoding huma
44	611.5	24.3	1803	24	Human kinase (PKIN
45	611.5	24.3	7626	24	Human ovary specif

# ALIGNMENTS

RESULT 1  
AAI60703  
ID AAI60703 standard; cDNA; 1956 BP.  
XX AAI60703;  
AC AAI60703;  
XX AAI60703;  
DT 22-OCT-2001 (first entry)  
XX Human polynucleotide SEQ ID NO 4692.  
DE  
DE  
XX Human; neoptotic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX  
XX Homo sapiens.  
OS



XX WO200153312-A1.  
 XX 26-JUL-2001.  
 XX 26-DEC-2000; 2000WO-US34263.  
 XX 21-JAN-2000; 2000US-0488725.  
 XX 25-APR-2000; 2000US-0552317.  
 XX 09-JUL-2000; 2000US-0598042.  
 XX 19-JUL-2000; 2000US-0620312.  
 XX 03-AUG-2000; 2000US-0653450.  
 XX 14-SEP-2000; 2000US-0652191.  
 XX 19-OCT-2000; 2000US-0693036.  
 XX 29-NOV-2000; 2000US-0727344.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX WPT; 2001-442253/47.  
 XX P-PSDB; AAM41547.  
 XX Novel nucleic acids and polypeptides, useful for treating disorders  
 XX such as central nervous system injuries -  
 XX Claim 1; SEQ ID NO 4692; 10078pp; English.  
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 XX immunosuppressant and cytostatic activity. The polynucleotides are useful  
 XX in gene therapy. A composition containing a polypeptide or polynucleotide  
 XX of the invention may be used to treat diseases of the peripheral nervous  
 XX system, such as peripheral nervous injuries, peripheral neuropathy and  
 XX localised neuropathies and central nervous system diseases, such as  
 XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 XX lateral sclerosis, and Shy-brager Syndrome. Other uses include the  
 XX utilisation of the activities such as: Immune system suppression,  
 XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 XX assays for receptor activity, arthritis and inflammation, leukaemias and  
 XX C.N.S disorders.  
 XX Note: The sequence data for this patent did not form part of the printed  
 XX specification.  
 XX SQ Sequence 1956 BP; 474 A; 559 C; 508 G; 414 T; 1 other;  
 Alignment Scores:  
 Pred. No.: 5.24e-206 Length: 1956  
 Score: 2513.00 Matches: 476  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-960-643-2 (1-476) x AAI60703 (1-1956)  
 QY 1 MetGlyArgLysGluGluAspAspCysSerSerTrpLysLysGlnThrThrAsnLeuArg 20  
 DB 65 ATGGGTCAAGAGGAGAGAGATGCTGAGTTCCTGGAAGAAACAGACCACCAATCCGG 124  
 QY 21 LysThrPheLeuPheMetGluValLeuGlySerGlyValAlaPheSerGluValPheLeuVal 40  
 DB 125 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGACCTTCTCAGAACTTTCTCGTG 184  
 QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysLysLysSerProAlaPhe 60  
 DB 185 AAGCAAGAACTGACTGGGAAGCTTTTGGCTCTGAAAGTGCATCAAGAAAGTCACTGCTTC 244  
 QY 61 ArgAspSerSerLeuGluAsnGluLeuAlaValLeuLysLysLysLysGlnAsnLeu 80  
 DB 1325 AGCTGCTGAACATTTGGGAGCAAGGAAAGTCTCTACTGCTGAGACCACTCTCTC 1384

DB 245 CGGGACAGCAGCCTGGAGAAATGAGATTGCTGTGTAAGAAAGATCAAGCATGAAACATT 304  
 QY 81 ValThrLeuGluAspPheLeuSerThrThrHisThrLeuValMetGlnLeuVal 100  
 DB 305 GTGACCTGGAGGACATCTATGAGAGCACCACCCTACTTACCTGGTCTATGACGCTTGT 364  
 QY 101 SerGlyGluLeuPheAspArgIleLeuGluArgGlyValThrGluLysAspAla 120  
 DB 365 TCTGTGGGGAGCTCTTTGACCGGATCTCGAGCGGGGTCTCTACACAGAGAGGATGCC 424  
 QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysThrLeuHisGluAsnGlyIleVal 140  
 DB 425 AGTCGTGGTATCCAGCAGGCTTGTGCGCAGTGAATACCTACTAGAAATGGCATCGTC 484  
 QY 141 HisArgAspLeuLysProGluAsnLeuLeuThrProGluGluAsnSerLysIle 160  
 DB 485 CACAGAGACTTAAAGCCCGAAACCTGCTTACCTTACCCTGAAGAGAACTCTAAGATC 544  
 QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180  
 DB 545 ATGATCACTGACTTGGTCTGCTCAAGATGGACAGAAATGGCATCTGCCACTGCTGT 604  
 QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200  
 DB 605 GGGACCCAGGCTAGCTGGCTCCAGAAAGTGTGCGCCAGAAACCTTACAGCAAGGCTGTG 664  
 QY 201 AspCysTyrSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProProPheTyr 220  
 DB 665 GATTGCTGGTCCATCGGGCTCATCACCTACATATTGCTGTGGATACCCCATTTCTAT 724  
 QY 221 GluGluThrGluSerLysLeuPheGluLysLysLysGluGlyTyrTyrGluPheGluSer 240  
 DB 725 GAAGAAGCGGATCTTAAAGCTTTTCGAGAAGATCAAGAGGGGCTACTATGATGTTGAGTCT 784  
 QY 241 ProPheTyrAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260  
 DB 785 CCATCTCTGGGATGACATTTCTGAGTCCAGCAGCAGGACCTTTATTTGCCACTTGTCTGAGAG 844  
 QY 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280  
 DB 845 GATCCGAAAGCGGCTACACCTGTGAGAAGGCCCTTGGATCTATCCCTGGATTCAGCGAAAC 904  
 QY 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300  
 DB 905 ACAGCCTCCACCGGAGACATCTACCATCAGCTCCAGCTCCAGATCCAGAGAAGACTTGTCT 964  
 QY 301 LysSerLysTyrArgGlnAlaPheAsnAlaAlaValHisHisMetArgLysLeu 320  
 DB 965 AAGAGCAAGTGGAGCAAGCCCTTCAAGCAGCAGCTGTGTGTCACCAACATCAGAGAACTA 1024  
 QY 321 HisMetAspLeuHisSerProGlyValArgProGluValGluAsnArgProGluThr 340  
 DB 1025 CACATGAACCTGCACAGCCCGGGGTCCGCCAGAGTGGAGAGACAGCCGCTGAAACT 1084  
 QY 341 GlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaProVal 360  
 DB 1085 CAAGCCTCAGAAACCTTAGACCCAGCTCCCTGAGATCACCATCAGCAGGACCTGTCT 1144  
 QY 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380  
 DB 1145 CTGGACACAGGTATGACACTCCCTGCGCTGACCCCAATTTACCTGCCAGCATGGCCCGG 1204  
 QY 381 ProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400  
 DB 1205 CCCACTGCCCTGGTGGCAGGTCCTCAACTGCCTGCTCAATGCTCCCTCCACATCAGC 1264  
 QY 401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420  
 DB 1265 AGCAGCCTGGTCCCATGCATCAGGGGTCCCTGGCCCGCCGGCCCTGTGCTGTCTCT 1324  
 QY 421 SerCysLeuAsnIleGlySerLysLysSerSerTyrCysSerGluProThrLeuLeu 440  
 DB 1325 AGCTGCTGAACATTTGGGAGCAAGGAAAGTCTCTACTGCTGAGACCACTCTCTC 1384



QY 441 LysLysAlaAsnLysLysGlnAsnPhelYsSerGluValMetValProValLysAlaSer 460  
 Db 1385 AAAAGGCCACAAAAACAGAACTTCAAGTCGGAGGTCAATGGTACCAGTTAAAGCCAGT 1444  
 QY 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476  
 Db 1445 GGCAGCTCCCACTGCCGGCAGGCAGACTGGAGTCTGTCATTTNG 1492

## RESULT 2

AAD36140  
 ID AAD36140 standard; DNA; 2447 BP.  
 XX AC  
 XX AAD36140;  
 XX 09-AUG-2002 (first entry)  
 XX DE Human calmodulin kinase, CAMK-X1 gene.  
 XX KW Human; cytostatic; antisense gene therapy; screening; protein kinase;  
 KW cancer; liver; colon; tumour; inflammation; arthritic synovium; CAMK-X1;  
 KW calmodulin kinase; enzyme; gene; chromosome 1q32.1-32.3; ds.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT CDS 70..1500  
 FT /\*tag= a  
 FT /product= "Human CAMK-X1 protein"  
 XX PN WO200224947-A2.  
 XX PD 28-MAR-2002.  
 XX PF 20-SEP-2001; 2001WO-IB02237.  
 XX PR 20-SEP-2000; 2000US-233999P.  
 PR 02-OCT-2000; 2000US-237419P.  
 PR 02-OCT-2000; 2000US-237423P.  
 PR 04-OCT-2000; 2000US-238558P.  
 PR 10-MAY-2001; 2001US-290555P.  
 XX (KINE-) KINETEK PHARM INC.  
 PA (UYBR-) UNIV BRITISH COLUMBIA.  
 XX Yoganathan T, Delaney AD;  
 XX WPI: 2002-394145/42.  
 DR P-PSDB; AAE22764.  
 XX Diagnosing cancer, comprises determining the upregulation of expression  
 of a nucleic acid sequence encoding a protein kinase or upregulation of  
 expression of the protein kinase, in the cancer  
 PS Claim 16; Page 62-64; 87pp; English.  
 XX The invention relates to a method for screening biologically active agent  
 CC that modulates cancer associated protein kinase function. The invention  
 CC also relates to a method for diagnosing cancer comprising determining the  
 CC upregulation of expression of a nucleic acid sequence encoding a protein  
 CC kinase. The method is useful for diagnosing cancer. A protein kinase is  
 CC useful for screening biological agents that modulate cancer associated  
 CC protein kinase function. Downregulating the activity of protein kinase is  
 CC useful for inhibiting the growth of a cancer cell, e.g. liver or colon  
 CC cancer. A nucleic acid encoding protein kinase is useful to screen biopsy  
 CC derived tumours and inflammatory samples such as arthritic synovium, for  
 CC amplified DNA in the cell or increased expression of corresponding mRNA  
 CC or protein and is also useful to detect differences in expression levels  
 CC such as molecular weight, amino acid and nucleotide sequences between the  
 CC two cells. The present sequence is human calmodulin kinase CAMK-X1 gene  
 CC located on chromosome 1q32.1-32.3.  
 XX SQ Sequence 2447 BP; 590 A; 707 C; 604 G; 546 T; 0 other;

## Alignment Scores:

Pred. No.: 7.2e-206 Length: 2447  
 Score: 2513.00 Matches: 476  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

US-09-960-643-2 (1-476) x AAD36140 (1-2447)

QY 1 MetGlyArgLysGluAspAspCysSerSerTyrLysLysGlnThrThrAsnIleArg 20  
 Db 70 ATGGGTGAAAGGAAGAGATGACTGCAGTTCTCTGGAAGAAACAGACCAACATCCGG 129  
 QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40  
 Db 130 AAAACCTTCATTTTATGGAGAGTCTGGGATCAGGAGCTTCTCAGAAGATTTTCCCTGGTG 189  
 QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60  
 Db 190 AAGCAAAACACTGACTGGGAGCTCTTTGCTCTGAAGTGCATCAAGAGTCACTGCCTTC 249  
 QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80  
 Db 250 CGGGACAGCAGCCTGGAGATGAGATTGCTGTGTTGAAAAGATCAAGCATGAAACAT 309  
 QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100  
 Db 310 GTGACCTGGAGGACATCTATGAGAGCACCACCTACTACTGTCATGAGAGCTGTT 369  
 QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120  
 Db 370 TCTGTGGGGAGCTCTTTGACCGGATCTGGAGCGGGTGTCTACACAGAGAAAGATGCC 429  
 QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140  
 Db 430 AGTCGTGGTGTCCAGCAGGTCTTCTCGCAGTGAATACCTCATGAGAAATGGCATCGTC 489  
 QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160  
 Db 490 CACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCCTGAAGAGAACTCTAAGATC 549  
 QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180  
 Db 550 ATGATCACTGACTTGTGTCTCCAGATGGACAGAAATGGCATCATGCTCCATGCTGT 609  
 QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200  
 Db 610 GGGACCCAGGCTAGTGGCTCCAGAAAGTGTGGCCCGAGAAACCTTACAGCAAGGCTGTG 669  
 QY 201 AspCysTyrSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProProPheTyr 220  
 Db 670 GATTCTGCTCCATCGGCTCATCCTCATATATGCTCTGTGGATACCCCGCTTCTAT 729  
 QY 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240  
 Db 730 GAAGAAACGGAGTCTTAAGCTTTTCGAGAAGATCAAGAGGGCTACTATGAGTTTGAGTCT 789  
 QY 241 ProPheTyrAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260  
 Db 790 CCATTCTGGGATGACATTTCTGAGTCAGCCCAAGGACTTTATTTGCCACTTCTTGAAGAG 849  
 QY 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProThrIleAspGlyAsn 280  
 Db 850 GATCGAAGCAGCGGTACACCTGTGAGAAGGCCCTTGAGTCATCCTCGATGATGACGGAAC 909  
 QY 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300  
 Db 910 ACGGCCCTCCACCGGAGATCATCTACCCATCAGTCAGCTCCAGATCCAGAGAACTTTGCT 969  
 QY 301 LysSerLysTyrArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu 320



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 10:03:49 ; Search time 566 Seconds  
(without alignments)  
9736.113 Million cell updates/sec

Title: US-09-960-643-1  
Perfect score: 2447  
Sequence: 1 tggagtggagctcaagcag.....ttttctctataaaaaaaa 2447

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				N_Geneseq_101002.*
1:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*			
2:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*			
3:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*			
4:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*			
5:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*			
6:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*			
7:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*			
8:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*			
9:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*			
10:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*			
11:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*			
12:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*			
13:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*			
14:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*			
15:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*			
16:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*			
17:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*			
18:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*			
19:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*			
20:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*			
21:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*			
22:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*			
23:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*			
24:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2447	100.0	2447	24	Human calmodulin k
2	1915	78.3	1956	22	Human polynucleoti
3	1854.6	75.8	2165	22	Human polynucleoti
4	1761.6	72.0	2689	22	Human diagnostic a
5	504.6	20.6	1565	24	Human Cam kinase I
6	504.6	20.6	1661	22	Human kinase (PKIN
7	504.6	20.6	1671	22	Nucleotide sequenc
8	504.6	20.6	1733	22	Human polynucleoti
9	503	20.6	1736	22	Human kinase (PKIN

10	503	20.6	2164	22	AA160424	Human polynucleoti
11	502.2	20.5	1074	22	AA506708	Polynucleotide seq
12	502.2	20.5	1074	22	AAH25119	Nucleotide sequenc
13	502.2	20.5	1158	22	AAH25118	Nucleotide sequenc
14	497	20.3	1578	22	AAAD0475	Human death domain
15	492.6	20.1	1694	22	AAAC90432	Murine Dendritic C
16	471.4	19.3	1728	21	AAZ46162	cDNA sequence enco
17	428.8	17.5	1458	19	AAV18867	Calmodulin-depende
18	427	17.4	1032	22	AA506711	Polynucleotide seq
19	426.2	17.4	588	23	ABK43480	DNA encoding novel
20	418	17.1	1282	20	AAK06835	Disease associated
21	393.8	16.1	1512	23	AA570157	DNA encoding novel
22	337.2	13.8	362	20	AAK51838	Human secreted pro
23	291	11.9	476	22	AA527154	cDNA encoding nove
24	291	11.9	476	23	ABK43797	DNA encoding novel
25	290	11.9	2264	23	ABL02719	Drosophila melanog
26	286.4	11.7	1572	22	AAH99632	Human protein enco
27	285.6	11.7	1774	23	AA568767	DNA encoding novel
28	285.6	11.7	1774	23	AA590919	DNA encoding novel
29	245.2	10.0	2897	22	AAK30481	Human protein phos
30	245.2	10.0	2905	22	AAK94325	Human full-length
31	245.2	10.0	3017	22	AAK18832	Human kinase (PKIN
32	243.6	10.0	2297	22	AAH78261	Nucleotide sequenc
33	243.6	10.0	2747	22	AAI58143	Human polynucleoti
34	243.6	10.0	2840	22	AAI58144	Human polynucleoti
35	243.6	10.0	3124	24	ABA95682	Human protein kina
36	242	9.9	2494	22	AAK94614	Human full-length
37	242	9.9	3022	22	AAK94299	DNA encoding novel
38	217.8	8.9	709	23	ABK43538	Human cDNA encodin
39	215.4	8.8	1181	22	AA534308	cDNA encoding nove
40	215.4	8.8	1181	22	AA527166	DNA encoding nove
41	215.4	8.8	1181	23	ABK43848	DNA encoding novel
42	197.2	8.1	3471	18	AAK71761	PSKH-1 cDNA. Homo
43	194	7.9	1386	22	AAK52586	Human polynucleoti
44	194	7.9	1450	22	AAK51602	Human polynucleoti
45	185.2	7.6	852	21	AAC76557	Human OREX ORF2112

ALIGNMENTS

RESULT 1  
AAD36140  
ID AAD36140 standard; DNA; 2447 BP.

AC AAD36140;  
XX  
XX 09-AUG-2002 (first entry)  
DT  
DE Human calmodulin kinase, CAMK-X1 gene.

XX Human; cytostatic; antisense gene therapy; screening; protein kinase;  
cancer; liver; colon; tumour; inflammation; arthritic synovium; CAMK-X1;  
KW calmodulin kinase; enzyme; gene; chromosome 1q32.1-32.3; ds.

OS Homo sapiens.  
XX Key Location/Qualifiers  
XX CDS 70..1500  
XX /\*tag= a  
XX /\*product= "Human CAMK-X1 protein"  
XX  
XX WO200224947-A2.

XX  
XX 28-MAR-2002.  
XX  
XX 20-SEP-2001; 2001WO-IB02237.  
XX  
XX 20-SEP-2000; 2000US-233999P.  
PR 02-OCT-2000; 2000US-237419P.  
PR 02-OCT-2000; 2000US-237423P.  
PR 04-OCT-2000; 2000US-238558P.  
PR 10-MAY-2001; 2001US-290555P.

XX (KINE-) KINETEK PHARM INC.  
PA (UYBR-) UNIV BRITISH COLUMBIA.  
XX  
PI Yoganathan T, Delaney AD;  
XX  
DR WPI; 2002-394145/42.  
DR P-PSDB; NAE22764.  
XX  
PT Diagnosing cancer, comprises determining the upregulation of expression  
PT of a nucleic acid sequence encoding a protein kinase or upregulation of  
PT expression of the protein kinase, in the cancer -  
XX  
PS Claim 16; Page 62-64; 87pp; English.  
XX  
CC The invention relates to a method for screening biologically active agent  
CC that modulates cancer associated protein kinase function. The invention  
CC also relates to a method for diagnosing cancer comprising determining the  
CC upregulation of expression of a nucleic acid sequence encoding a protein  
CC kinase. The method is useful for diagnosing cancer. A protein kinase is  
CC useful for screening biological agents that modulate cancer associated  
CC protein kinase function. Downregulating the activity of protein kinase is  
CC useful for inhibiting the growth of a cancer cell, e.g. liver or colon  
CC cancer. A nucleic acid encoding protein kinase is useful to screen biopsy  
CC derived tumours and inflammatory samples such as arthritic synovium, for  
CC amplified DNA in the cell or increased expression of corresponding mRNA  
CC or protein and is also useful to detect differences in expression levels  
CC such as molecular weight, amino acid and nucleotide sequences between the  
CC two cells. The present sequence is human calmodulin kinase CAMK-X1 gene  
CC located on chromosome 1q32.1-32.3.  
XX  
SQ Sequence 2447 BP; 590 A; 707 C; 604 G; 546 T; 0 other;

Query Match  
Best Local Similarity 100.0%; Score 2447; DB 24; Length 2447;  
Matches 2447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGAGTGGAGCTCAAGCAGGATTCTCCGAGTCCCTGGCATCCTCAGAAAGTTCAACT 60  
Db 1 TGGAGTGGAGCTCAAGCAGGATTCTCCGAGTCCCTGGCATCCTCAGAAAGTTCAACT 60

QY 61 CTGGAGGCAATGGGTGCGAAGGAAGATGACTGCACTTCTTGGAGCAAGCAACACC 120  
Db 61 CTGGAGGCAATGGGTGCGAAGGAAGATGACTGCACTTCTTGGAGCAAGCAACACC 120

QY 121 AACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTCTCAGAAAGTT 180  
Db 121 AACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTCTCAGAAAGTT 180

QY 181 TTCCTGGTGAAGCAAGACTGACTGGGAAGTCTTTGCTGTAAGTGCATCAAGAAGTCA 240  
Db 181 TTCCTGGTGAAGCAAGACTGACTGGGAAGTCTTTGCTGTAAGTGCATCAAGAAGTCA 240

QY 241 CTTGCTTCCGGGACAGCCTCGAGATGAGTTGCTGTTGGAAGCAAGCAAGCAT 300  
Db 241 CTTGCTTCCGGGACAGCCTCGAGATGAGTTGCTGTTGGAAGCAAGCAAGCAT 300

QY 301 GAAACATTTGACCTCGAGGACATCTATGAGACACACCACCTACTACTGTCATG 360  
Db 301 GAAACATTTGACCTCGAGGACATCTATGAGACACACCACCTACTACTGTCATG 360

QY 361 CAGCTTTGTTCTGGTGGGAGTCTTTGACCGGATCCTTGGAGCGGGGTGTCTACACAG 420  
Db 361 CAGCTTTGTTCTGGTGGGAGTCTTTGACCGGATCCTTGGAGCGGGGTGTCTACACAG 420

QY 421 AAGGATGCCAGTCTGGTATCCAGAGTCTTTGTCGGCAGTGAATACCTACATGAGAA 480  
Db 421 AAGGATGCCAGTCTGGTATCCAGAGTCTTTGTCGGCAGTGAATACCTACATGAGAA 480

QY 481 GGCATCGTCCACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC 540  
Db 481 GGCATCGTCCACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC 540

QY 541 TCTAAGATCATGATGACTGACTTTGGTCTGTCCAAAGATGGAACAGAAATGCGATCATGTCC 600  
Db 541 TCTAAGATCATGATGACTGACTTTGGTCTGTCCAAAGATGGAACAGAAATGCGATCATGTCC 600

QY 601 ACTGCCCTGTGGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCAGAAACCTTACAGC 660  
Db 601 ACTGCCCTGTGGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCAGAAACCTTACAGC 660

QY 661 AAGGCTGTGATGTGCTGCTCCATCGGCGTCACTCACCTACATATTGCTGTGTGATACCCC 720  
Db 661 AAGGCTGTGATGTGCTGCTCCATCGGCGTCACTCACCTACATATTGCTGTGTGATACCCC 720

QY 721 CCGTCTTATGAAGAACGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAG 780  
Db 721 CCGTCTTATGAAGAACGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAG 780

QY 781 TTTGAGTCTCCATTTCTGGGATGACATTTCTGAGTACGCCAAGGACTTTATTATGCCACTTG 840  
Db 781 TTTGAGTCTCCATTTCTGGGATGACATTTCTGAGTACGCCAAGGACTTTATTATGCCACTTG 840

QY 841 CTTGAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGGCTTTGAGTCAATCCCTGGATT 900  
Db 841 CTTGAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGGCTTTGAGTCAATCCCTGGATT 900

QY 901 GAGGAAACAGGCGCTCCACCGGGACATCTACCCATCAGTCAAGCTCCAGATCCAGAAAG 960  
Db 901 GAGGAAACAGGCGCTCCACCGGGACATCTACCCATCAGTCAAGCTCCAGATCCAGAAAG 960

QY 961 AACTTTCTTAAGAGCAAGTGGAGCAAGCTTCAACGACAGCAGTGTGGTGCACCAATG 1020  
Db 961 AACTTTCTTAAGAGCAAGTGGAGCAAGCTTCAACGACAGCAGTGTGGTGCACCAATG 1020

QY 1021 AGGAAGCTACACATGAACCTGCACAGCCCGGGGTCCGCCAGAGGTGGAGAAACAGGCCG 1080  
Db 1021 AGGAAGCTACACATGAACCTGCACAGCCCGGGGTCCGCCAGAGGTGGAGAAACAGGCCG 1080

QY 1081 CTTGAAACTCAAGCTCAGAAACCTCTAGACCAGCTTCCCTGAGATCAACATCAACCGAG 1140  
Db 1081 CTTGAAACTCAAGCTCAGAAACCTCTAGACCAGCTTCCCTGAGATCAACATCAACCGAG 1140

QY 1141 GCACCTCTCTGGACACAGTGTAGCAGTCCCTGACCAATTTACCTGCGCAGCAT 1200  
Db 1141 GCACCTCTCTGGACACAGTGTAGCAGTCCCTGACCAATTTACCTGCGCAGCAT 1200

QY 1201 GGCGCGCGGCCACTGCCCCCTGCTGAGTGTCCCTCAACTGCTGCTCAATGCTCCCTC 1260  
Db 1201 GGCGCGCGGCCACTGCCCCCTGCTGAGTGTCCCTCAACTGCTGCTCAATGCTCCCTC 1260

QY 1261 CACATCAGCAGCAGCTGCTGCCCCATGCATCAGGGGTCCCTGGCGCGCGGCCCTGTGGC 1320  
Db 1261 CACATCAGCAGCAGCTGCTGCCCCATGCATCAGGGGTCCCTGGCGCGCGGCCCTGTGGC 1320

QY 1321 TGCTGCTCCAGCTGCTGAACATTTGGGAGCAAGAAAGTCCCTACTGCTCTGAGCCCC 1380  
Db 1321 TGCTGCTCCAGCTGCTGAACATTTGGGAGCAAGAAAGTCCCTACTGCTCTGAGCCCC 1380

QY 1381 ACATCTCTAAAAGGCGCAACAAAACAGAACTTCAAGTGGAGGTCAATGGTACAGATT 1440  
Db 1381 ACATCTCTAAAAGGCGCAACAAAACAGAACTTCAAGTGGAGGTCAATGGTACAGATT 1440

QY 1441 AAAGCCAGTGGAGTCCCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500  
Db 1441 AAAGCCAGTGGAGTCCCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500

QY 1501 TTCTGAGGAGCTGTGCTGCTATGCTCAATTTTCAGGAGACATATTCAACTCCTCTGCT 1560  
Db 1501 TTCTGAGGAGCTGTGCTGCTATGCTCAATTTTCAGGAGACATATTCAACTCCTCTGCT 1560

QY 1561 CTTCCAAACCTGTGTCTATCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGCTT 1620  
Db 1561 CTTCCAAACCTGTGTCTATCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGCTT 1620

QY 1621 AGCAGGAGGAGTCTTCTGGCCAGAAAGCAGCAGCTGCTGCCAGCGGGGAGCCCTCATAG 1680

Db 1621 AGCAGGAGAGTTCTTGGCCAGAGACACAGCCTGTGTCAGGGGGGCGCCCTCATAG 1680  
Qy 1681 GAGGCCAGGAGGAGCCCAAGCGCTAGAGCCTTGTGTTGAAGCTGTGAGCAGGAGAAC 1740  
Db 1681 GAGGCCAGGAGGAGCCCAAGCGCTAGAGCCTTGTGTTGAAGCTGTGAGCAGGAGAAC 1740  
Qy 1741 GGTGCCACACAGCTTCCAGTCTCCCTGACCTGCCTGCTCTATGCCCCACACCCCTACGTG 1800  
Db 1741 GGTGCCACACAGCTTCCAGTCTCCCTGACCTGCCTGCTCTATGCCCCACACCCCTACGTG 1800  
Qy 1801 CCGTGGCTCTGTGCACTGACGTAGTAGTCTCGGCTGGGTGTGCTGCTGTTTGTCTGTA 1860  
Db 1801 CCGTGGCTCTGTGCACTGACGTAGTAGTCTCGGCTGGGTGTGCTGCTGTTTGTCTGTA 1860  
Qy 1861 AAAGCTTAATGGGTGGCCAGGTGTGTACCTTCTCCAGCAAGCCCATATGAGCATC 1920  
Db 1861 AAAGCTTAATGGGTGGCCAGGTGTGTACCTTCTCCAGCAAGCCCATATGAGCATC 1920  
Qy 1921 TACCCAGACTCCCACTCTGCACACACTCACTCCCACTCTCAAGCCTCCCACTCTTGGC 1980  
Db 1921 TACCCAGACTCCCACTCTGCACACACTCACTCCCACTCTCAAGCCTCCCACTCTTGGC 1980  
Qy 1981 CAGATTGGGTCAATTAATGTGCTGCTGCCATCTGCAATGAATGACAGCAGCTCCCA 2040  
Db 1981 CAGATTGGGTCAATTAATGTGCTGCTGCCATCTGCAATGAATGACAGCAGCTCCCA 2040  
Qy 2041 TGGTGTCTGCTGTGAGCTTCTCAAGTTCTTAATCTTAATCTTAATCTTAATCTTAAT 2100  
Db 2041 TGGTGTCTGCTGTGAGCTTCTCAAGTTCTTAATCTTAATCTTAATCTTAATCTTAAT 2100  
Qy 2101 TGGCTGAGACCCAGCAGCAGCAGCTTCTGCGCTCTCTGCTGCTCAATCTAAAGCAGT 2160  
Db 2101 TGGCTGAGACCCAGCAGCAGCAGCTTCTGCGCTCTCTGCTGCTCAATCTAAAGCAGT 2160  
Qy 2161 GCCACACCCCTCCAAAGTGAATGAAGAAGTTTCATGAGTAAGGGCTGCAAGGAATCTT 2220  
Db 2161 GCCACACCCCTCCAAAGTGAATGAAGAAGTTTCATGAGTAAGGGCTGCAAGGAATCTT 2220  
Qy 2221 ATCTGCGCCACATCTCTCCGTCGACACACCCCAATGAGTTACCTTGGAGTTGACTAT 2280  
Db 2221 ATCTGCGCCACATCTCTCCGTCGACACACCCCAATGAGTTACCTTGGAGTTGACTAT 2280  
Qy 2281 TTTAATGTCTGCCAGGAGTTCTAATCTGCTCTGTTTCTCTCTCTCTCTCTCTCTCTCT 2340  
Db 2281 TTTAATGTCTGCCAGGAGTTCTAATCTGCTCTGTTTCTCTCTCTCTCTCTCTCTCTCT 2340  
Qy 2341 AGCACACATCTTGTCTCCGAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400  
Db 2341 AGCACACATCTTGTCTCCGAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400  
Qy 2401 AGTGTGTGCTTAATAAATGGACATATTTTCTCTAAAGAAAAA 2447  
Db 2401 AGTGTGTGCTTAATAAATGGACATATTTTCTCTAAAGAAAAA 2447

RESULT 2  
AAI60703  
ID AAI60703 standard; cdna; 1956 BP.  
XX  
AC AAI60703;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 4692.  
XX  
KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.

OS Homo sapiens.  
PN WO200153312-A1.  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US34263.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI: 2001-442253/47.  
XX P-PSDB; AAM41547.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
PS Claim 1; SEQ ID NO 4692; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nontropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.

SQ Sequence 1956 BP; 474 A; 559 C; 508 G; 414 T; 1 other;  
Query Match 78.3%; Score 1915; DB 22; Length 1956;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1941; Conservative 0; Mismatches 6; Indels 3; Gaps 2;  
Qy 12 CTCAAGCAGGATTTCTCCCGAGTCCCTGGCATCTCTCAGAAGCTTCAACTCTGGAGCAAT 71  
Db 7 CTCAAGCAGGATTTCTCCCGAGTCCCTGGCATCTCTCAGAAGCTTCAACTCTGGAGCAAT 66  
Qy 72 GGGTCGAAGAGAGATGACTGCAGTTCTCTGGAGAAACAGACCACACATCCGGAA 131  
Db 67 GGGTCGAAGAGAGATGACTGCAGTTCTCTGGAGAAACAGACCACACATCCGGAA 126  
Qy 132 AACCTTCATTTTATGGAAGTGGGATCAGGAGCTTCTCAGAAGTTTCTCGGTGAA 191  
Db 127 AACCTTCATTTTATGGAAGTGGGATCAGGAGCTTCTCAGAAGTTTCTCGGTGAA 186  
Qy 192 GCAGAGACTGAGTGGGAAGCTTTTGTCTCTGAAGTGCATCAAGAGTCACCTGCTCCG 251  
Db 187 GCAGAGACTGAGTGGGAAGCTTTTGTCTCTGAAGTGCATCAAGAGTCACCTGCTCCG 246  
Qy 252 GCAGAGAGCTGGGAAGATGAGTTGCTGTGTTGAAAAGATCAAGCATGAAACATGTT 311  
Db 247 GCAGAGAGCTGGGAAGATGAGTTGCTGTGTTGAAAAGATCAAGCATGAAACATGTT 306

QY 312 GACCCCTGGAGGACATCTATGAGAGACACACCACCTACTACCTGGTTCATGCGAGCTGTGTTTC 371  
DB 307 GACCCCTGGAGGACATCTATGAGAGACACACCACCTACTACCTGGTTCATGCGAGCTGTGTTTC 366  
QY 372 TGGTGGGAGCTCTTTGACCGGATCCTGGAGGGGTGTCTACACAGAGAAAGATGCCAG 431  
DB 367 TGGTGGGAGCTCTTTGACCGGATCCTGGAGGGGTGTCTACACAGAGAAAGATGCCAG 426  
QY 432 TCTGGTGATCCAGCAGGCTTGTGGCAGTGAAATACCTACATGAGATGGCATCGTCCA 491  
DB 427 TCTGGTGATCCAGCAGGCTTGTGGCAGTGAAATACCTACATGAGATGGCATCGTCCA 486  
QY 492 CAGAGACTTTAAAGCCGGAACCTGCTTTACCTTACCCCTGAAGAGAACTCTTAAGATCAT 551  
DB 487 CAGAGACTTTAAAGCCGGAACCTGCTTTACCTTACCCCTGAAGAGAACTCTTAAGATCAT 546  
QY 552 GATCACTGACTTTGGTCTCTCCAAGATGGAACAGAAATGGCATCATGTCCACTGCGCTGRGG 611  
DB 547 GATCACTGACTTTGGTCTCTCCAAGATGGAACAGAAATGGCATCATGTCCACTGCGCTGRGG 606  
QY 612 GACCCAGGCTAGCTGGCTCCAGAGTGTGGCCAGAAACCTACAGCAAGGCTGTGGA 671  
DB 607 GACCCAGGCTAGCTGGCTCCAGAGTGTGGCCAGAAACCTACAGCAAGGCTGTGGA 666  
QY 672 TTGCTGGTCCATCGGCGTCACTACCTACATATTGCTCTGTGATACCCCGCTTCTATGA 731  
DB 667 TTGCTGGTCCATCGGCGTCACTACCTACATATTGCTCTGTGATACCCCGCTTCTATGA 726  
QY 732 AGAAACGGAGTCAAGCTTTTCAGAGATCAAGAGGGCTACTATGATTTGAGTCTCC 791  
DB 727 AGAAACGGAGTCAAGCTTTTCAGAGATCAAGAGGGCTACTATGATTTGAGTCTCC 786  
QY 792 ATTCTGGGATGACATTTCTGAGTCAAGCAAGGACTTTATTTCGCACTGCTTTGAGAAGA 851  
DB 787 ATTCTGGGATGACATTTCTGAGTCAAGCAAGGACTTTATTTCGCACTGCTTTGAGAAGA 846  
QY 852 TCCGAACGAGCGGTACACCTGTGAGAGGCCCTTGAGTCAATCCCTGGATGACGGAACAC 911  
DB 847 TCCGAACGAGCGGTACACCTGTGAGAGGCCCTTGAGTCAATCCCTGGATGACGGAACAC 906  
QY 912 GGCCCTCCACCGGAGATCTACCCATCAGTCAGCTCCAGTCCAGAGTCCAGAGAACTTTGCTAA 971  
DB 907 AGCCCTCCACCGGAGATCTACCCATCAGTCAGCTCCAGTCCAGAGAACTTTGCTAA 966  
QY 972 GAGCAAGTGGAGCAAGCCTTCAACGACGACGCTGTGGTGCACCATGAGGAAGCTACA 1031  
DB 967 GAGCAAGTGGAGCAAGCCTTCAACGACGACGCTGTGGTGCACCATGAGGAAGCTACA 1026  
QY 1032 CATGAACCTGCACAGCCCGGGGTCCGCCAGAGGTGGAGAACAGCCGCTGAAACTCA 1091  
DB 1027 CATGAACCTGCACAGCCCGGGGTCCGCCAGAGGTGGAGAACAGCCGCTGAAACTCA 1086  
QY 1092 AGCCTCAGAAACCTCTAGACCCAGCTCCCTTGAGATCACCATCACCGAGGACCTGTCT 1151  
DB 1087 AGCCTCAGAAACCTCTAGACCCAGCTCCCTTGAGATCACCATCACCGAGGACCTGTCT 1146  
QY 1152 GGACCACAGTGTAGCACTCCCTGCCCTGACCCAAATPACCTGCCAGATGGCCGCGGCGC 1211  
DB 1147 GGACCACAGTGTAGCACTCCCTGCCCTGACCCAAATPACCTGCCAGATGGCCGCGGCGC 1206  
QY 1212 CACTGCCCTGGTGGCAGGTCCCTCAACTGCCCTGGTGGTCAATGGCTCCCTCCACATCAGCAG 1271  
DB 1207 CACTGCCCTGGTGGCAGGTCCCTCAACTGCCCTGGTGGTCAATGGCTCCCTCCACATCAGCAG 1266  
QY 1272 CAGCTGGTGGTCCCATGCATCAGGGGTCCCTGGCCGCGGGGCTGTGGCTGTGCTGCCAG 1331  
DB 1267 CAGCTGGTGGTCCCATGCATCAGGGGTCCCTGGCCGCGGGGCTGTGGCTGTGCTCCAG 1326  
QY 1332 CTGCCCTGAACATTTGGGAGCAAAAGTCTCTCTACTGCTGTGAGCCACACTCTCTCAA 1391  
DB 1327 CTGCCCTGAACATTTGGGAGCAAAAGTCTCTCTACTGCTGTGAGCCACACTCTCTCAA 1386

QY 1392 AAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTCAATGGTACCACTTAAAGCCAGTGG 1451  
DB 1387 AAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTCAATGGTACCACTTAAAGCCAGTGG 1446  
QY 1452 CAGCTCCCACTCCCGGCGGAGAGCTGGAGTCTGTCTCATTTATGTGATTCTCTGGAGCC 1511  
DB 1447 CAGCTCCCACTCCCGGCGGAGAGCTGGAGTCTGTCTCATTTATGTGATTCTCTGGAGCC 1506  
QY 1512 TGTGCTTATGTCACTGCAATTTTTCAGGAGACATATTCAACTCCCTCTCTTCCAAACCT 1571  
DB 1507 TGTGCTTATGTCACTGCAATTTTTCAGGAGACATATTCAACTCCCTCTCTTCCAAACCT 1566  
QY 1572 GGTGCTTATCCGCGAGAGGAGGAGAGCAGCAAGTGGAGCAGGGCTTAGCAGAGCAG 1631  
DB 1567 GGTGCTTATCCGCGAGAGGAGGAGAGCAGCAAGTGGAGCAGGGCTTAGCAGAGCAG 1626  
QY 1632 TTTCTGGCCAGAACACCAGCCTGCTGCCAGCGGGGAGCCCTCATAGGAGGCCCAGGA 1691  
DB 1627 TTTCTGGCCAGAACACCAGCCTGCTGCCAGCGGGGAGCCCTCATAGGAGGCCCAGGA 1686  
QY 1692 GGGAGCCCCAAGCGGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAAAGCGGTGCCACCA 1751  
DB 1687 GGGAGCCCCAAGCGGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAAAGCGGTGCCACCA 1746  
QY 1752 GCTTCCAGGTCTCCCTGACCTGCTGTATGATGCCACACACCTTACGT-GCCGTGGCTCT 1810  
DB 1747 GCTTCCAGGTCTCCCTGACCTGCTGTATGATGCCACACACCTTACGTGGCCGTGGCTCT 1806  
QY 1811 GTGCAGTGT--ACCTAGATAGCTCTCGCTGGGTGTGCTGTGTTGCTGAAAAGCTTA 1868  
DB 1807 GTGCAGTGTACGTTAGATAGCTCTCGCTGGGTGTGCTGTGTTGCTGAAAAGCTTA 1866  
QY 1869 ATGGCTGGCCAGGCTGTGTCACTTCTTCCAAGCAAGCCATATGGAGCATCTACCCAGA 1928  
DB 1867 ATGGCTGGCCAGGCTGTGTCACTTCTTCCAAGCAAGCCATATGGAGCATCTACCCAGA 1926  
QY 1929 CTCACACTCTGCACACACTCACTCCACCT 1958  
DB 1927 CTCACACTCTGCACACACTCACTCCACCT 1956

RESULT 3

AAI58917  
ID AAI58917 standard; cDNA; 2165 BP.  
XX AAI58917;  
XX AC  
XX AC  
DT 22-Oct-2001 (first entry)  
XX Human polynucleotide SEQ ID NO 1120.  
DE  
XX  
KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
OS Homo sapiens.  
XX  
XX WO200153312-A1.  
XX  
XX 26-JUL-2001.  
PD  
XX 26-DEC-2000; 2000WO-US34263.  
PF  
XX 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.

29-NOV-2000; 2000US-0727344.  
(HYSE-) HYSEQ INC.  
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
WPI; 2001-442253/47.  
P-PSDB; AAM39761.  
Novel nucleic acids and polypeptides, useful for treating disorders  
such as central nervous system injuries -  
Claim 1; SEQ ID NO 1120; 10078pp; English.  
The invention relates to human nucleic acids (AA157798-AA161369) and  
the encoded polypeptides (AAM38642-AA42213) with nootropic,  
immunosuppressant and cytoskeletal activity. The polynucleotides are useful  
in gene therapy. A composition containing a polypeptide or polynucleotide  
of the invention may be used to treat diseases of the peripheral nervous  
system, such as peripheral nervous injuries, peripheral neuropathy and  
localised neuropathies and central nervous system diseases, such as  
Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
utilisation of the activities such as: Immune system suppression,  
Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
assays for receptor activity, arthritis and inflammation, leukaemias and  
C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
specification.  
Query Match 75.8%; Score 1854.6; DB 22; Length 2165;  
Best Local Similarity 98.3%; Pred. No. 0;  
Matches 1914; Conservative 0; Mismatches 10; Indels 24; Gaps 3;  
35 CCCTGGGATCCTCAGAGCTTCAACCTGGAGGCAATGGTCAAGAGGAAGATGACT 94  
218 CACACCGAGCTCCTCAGAGCTTCAACCTGGAGGCAATGGTCAAGAGGAAGATGACT 277  
95 GCAGTTCCTGGAAGAAACACACCAACATCCGGAACCTTCATTTTATGGAAGTGC 154  
278 GCAGTTCCTGGAAGAAACACACCAACATCCGGAACCTTCATTTTATGGAAGTGC 337  
155 TGGGATCAGGAGCTTCTCAGAAAGTTTCTCGTGAAGCAAAAGACTGACTGGGAAGCTCT 214  
338 TGGGATCAGGAGCTTCTCAGAAAGTTTCTCGTGAAGCAAAAGACTGACTGGGAAGCTCT 397  
215 TTGCTCTGAAGTGCATCAAGATCAGCTCCCTTCGGGACAGCAGCCTGGAGATGAGA 274  
398 TTGCTCTGAAGTGCATCAAGATCAGCTCCCTTCGGGACAGCAGCCTGGAGATGAGA 457  
275 TTGCTGTGTTGAAAGATCAAGCATGAAACATTTGACCCCTGGAGGACATCTATGAGA 334  
458 TTGCTGTGTTGAAAGATCAAGCATGAAACATTTGACCCCTGGAGGACATCTATGAGA 517  
335 GCACCAACCCACTACTACCTGGTGCATGAGCTTGTCTTCTGGTGGGAGCTCTTTGACCGGA 394  
518 GCACCAACCCACTACTACCTGGTGCATGAGCTTGTCTTCTGGTGGGAGCTCTTTGACCGGA 577  
395 TCCTGGAGCGGGTGTCTACACAGAGAGATGCCAGTCTGGTGATCCAGAGGCTTGT 454  
578 TCCTGGAGCGGGTGTCTACACAGAGAGATGCCAGTCTGGTGATCCAGAGGCTTGT 637  
455 CGGCAGTGAATACCTACATGAGATGGCATCGTCACAGAGACTTAAA----- 503  
638 CGGCAGTGAATACCTACATGAGATGGCATCGTCACAGAGACTTAAAAGGTGTCAGGC 697  
504 -----GCCCGAAACCTGCTTTACCTTACCCTTGAAGAGAACTCTTAAGATCATGA 553  
|||||

Db 698 GGGAGTCTCTGGCCGCAAAACCTGCTTTACCTTACCCCTGGAAGAGAACTCTAAGATCATGA 757  
QY 554 TCACGTGACTTTGGTCTGTCCAAAGATGGAACAGAAATGGCATCATGTCCACTGCTGTGGGA 613  
Db 758 TCACGTGACTTTGGTCTGTCCAAAGATGGAACAGAAATGGCATCATGTCCACTGCTGTGGGA 817  
QY 614 CCCAGGCTACGTGCTCCAGAGTGTCTGGCCAGAAACCTTACAGAAAGCTCTGTGATTT 673  
Db 818 CCCAGGATACGTGGCTCCAGAGTGTCTGGCCAGAAACCTTACAGAAAGCTCTGTGATTT 877  
QY 674 GCTGTCCATCGGCTCATACCTACATATTGCTCTGTGGATACCCCGCTTCTATGAAG 733  
Db 878 GCTGTCCATCGGCTCATACCTACATATTGCTCTGTGGATACCCCGCTTCTATGAAG 937  
QY 734 AAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGGCTACTATGAGTTTGAGTCTCCAT 793  
Db 938 AAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGGCTACTATGAGTTTGAGTCTCCAT 997  
QY 794 TCTGGATGACATTTCTCAGTCAGCAAGGACTTTATTGTCACCTTGCTTGAGAAAGATC 853  
Db 998 TCTGGATGACATTTCTCAGTCAGCAAGGACTTTATTGTCACCTTGCTTGAGAAAGATC 1057  
QY 854 CGAAGAGCGGTACACCTGTGAGAAAGGCTTTGAGTCACTCCCTGGATTGACGAAACACGG 913  
Db 1058 CGAAGAGCGGTACACCTGTGAGAAAGGCTTTGAGTCACTCCCTGGATTGACGAAACACAG 1117  
QY 914 CCTCCACCGGGACATCTACCCATCAGTCAGCTCCAGATCCAGAAAGCTTTGCTTGAAGA 973  
Db 1118 CCTCCACCGGGACATCTACCCATCAGTCAGCTCCAGATCCAGAAAGCTTTGCTTGAAGA 1177  
QY 974 GCAAGTGAGGAGCAAGCTTTCAAGCAGCAGCTGTGGTGACACACATGAGGAAGCTACACA 1033  
Db 1178 GCAAGTGAGGAGCAAGCTTTCAAGCAGCAGCTGTGGTGACACACATGAGGAAGCTACACA 1237  
QY 1034 TGAACCTGCACAGCCCGGGCTCCGCCAGAGGTGGAGAACAGCCGCTGAAACTCAAG 1093  
Db 1238 TGAACCTGCACAGCCCGGGCTCCGCCAGAGGTGGAGAACAGCCGCTGAAACTCAAG 1297  
QY 1094 CCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCACCATCCAGGAGCAGCTGCTCTGG 1153  
Db 1298 COTCAGAAACCTCTAGACCCAGCTCCCTGAGATCACCATCCAGGAGCAGCTGCTCTGG 1357  
QY 1154 ACCACAGTGTAGCAGCTCCCTGACCCCAATTTACCTGCGCAGCATGGCCGCGCGGCCA 1213  
Db 1358 ACCACAGTGTAGCAGCTCCCTGCGCTGACCCCAATTTACCTGCGCAGCATGGCCGCGGCCA 1417  
QY 1214 CTGCCCCCTGGTGGCAGGTCCCTCAACTGCTTGGTCAATGGCTCCCTCCACATCAGCAGCA 1273  
Db 1418 CTGCCCCCTGGTGGCAGGTCCCTCAACTGCTTGGTCAATGGCTCCCTCCACATCAGCAGCA 1477  
QY 1274 GCCTGGTGCCCATGCATCAGGGGTCCCTGGCCGCGGGGCTGTGGCTGCTGCTCCAGCT 1333  
Db 1478 GCCTGGTGCCCATGCATCAGGGGTCCCTGGCCGCGGGGCTGTGGCTGCTGCTCCAGCT 1537  
QY 1334 GCCTGAACATTTGGGAGCAAGAAAGTCTCTACTGCTGTGAGCCACACACTCCCTCAAAA 1393  
Db 1538 GCCTGAACATTTGGGAGCAAGAAAGTCTCTACTGCTGTGAGCCACACACTCCCTCAAAA 1597  
QY 1394 AGGCCAACAAAAACAGAACTTCAAGTCGAGGTCATGGTACCAGTTAAAGCCAGTGGCA 1453  
Db 1598 AGGCCAACAAAAACAGAACTTCAAGTCGAGGTCATGGTACCAGTTAAAGCCAGTGGCA 1657  
QY 1454 GCTCCCACTGCGGCGCAGGCGAGCTGGAGTCTGTCTCATTTATGATTTCTGAGGCGCTG 1513  
Db 1658 GCTCCCACTGCGGCGCAGGCGAGCTGGAGTCTGTCTCATTTATGATTTCTGAGGCGCTG 1717  
QY 1514 TGCCCTATGCTCACTGCAATTTTCAGGAGACATATTCAACTCCTCTGCTTCTCCAAACCTGG 1573  
Db 1718 TGCCCTATGCTCACTGCAATTTTCAGGAGACATATTCAACTCCTCTGCTTCTCCAAACCTGG 1777  
QY 1574 TGCTCTATCCGCGAGAGGAGGAGGAGCAGAGCAAGTGGAGCAGGGCTTTAGCAGGAGCAGTT 1633  
Db 1778 TGCTCTATCCGCGAGAGGAGGAGGAGGAGCAGAGCAAGTGGAGCAGGGCTTTAGCAGGAGCAGTT 1837



Db 386 TCTGAAGTCATCAAGAAGTCACTGCCTTCGGGACAGCAGCCTGGAGAATGAGATTGC 445  
QY 279 TGTGTTGAAAAGATCAAGCATGAACAACATTTGTGACCCCTGGAGGACATCTATGAGAGCAC 338  
Db 446 TGTGTTGAAAAGATCAAGCATGAACAACATTTGTGACCCCTGGAGGACATCTATGAGAGCAC 505  
QY 339 CACCCACTACTACTCGTTCATGCAGCTTGTCTTGTGGGGAGCTCTTTGACCGGATCCT 398  
Db 506 CACCCACTACTACTCGTTCATGCAGCTTGTCTTGTGGGGAGCTCTTTGACCGGATCCT 565  
QY 399 GGAGCGGGTGTCTACACAGAGAAGGATGCCAGTCTGTGTGATCCACAGAGCTTAAAGCCGGAAAACCTGCT 458  
Db 566 GGAGCGGGTGTCTACACAGAGAAGGATGCCAGTCTGTGTGATCCACAGAGCTTAAAGCCGGAAAACCTGCT 625  
QY 459 AGTGAATACCTACATGAGAATGGCATCTAAGATCATGATCACTGACTTTGGTCTCTCCAAGAT 518  
Db 626 AGTGAATACCTACATGAGAATGGCATCTAAGATCATGATCACTGACTTTGGTCTCTCCAAGAT 685  
QY 519 TTACCTTACCCCTGAAGAGAACCTCTAAGATCATGATCACTGACTTTGGTCTCTCCAAGAT 578  
Db 686 TTACCTTACCCCTGAAGAGAACCTCTAAGATCATGATCACTGACTTTGGTCTCTCCAAGAT 745  
QY 579 GGAACGAATGGCATCATGTCCACTGCCTGTGGGACCCAGGCTACGTGGCTCCAGAACT 638  
Db 746 GGAACGAATGGCATCATGTCCACTGCCTGTGGGACCCAGGCTACGTGGCTCCAGAACT 805  
QY 639 GCTGGCCCAAAACCTACAGAAAGCTGTGATTTGCTGGTCCATCGGGGTCAACCTA 698  
Db 806 GCTGGCCCAAAACCTACAGAAAGCTGTGATTTGCTGGTCCATCGGGGTCAACCTA 865  
QY 699 CATATTGCTCTGTGGATACCCCGCTTCTATGAAGAAACGGAGTCTAAGCTTTTCAGAAA 758  
Db 866 CATATTGCTCTGTGGATACCCCGCTTCTATGAAGAAACGGAGTCTAAGCTTTTCAGAAA 925  
QY 759 GATCAAGAGGGCTACTATGAGTTTGTGATCTCCATTTCTGGGATGACATTTCTGAGTCAGC 818  
Db 926 GATCAAGAGGGCTACTATGAGTTTGTGATCTCCATTTCTGGGATGACATTTCTGAGTCAGC 985  
QY 819 CAAGGACTTTATTGTCACCTGCTTGAGAGGATCCGAACGATCCGAACGA - GGGGTACACCTGTGAGA 877  
Db 986 CAAGGACTTTATTGTCACCTGCTTGAGAGGATCCGAACGAGGGGTAAACACTGTGAGA 1045  
QY 878 AGGCCTTGAGTCATCCCTGGATTGACGGAACACGSCCCTCCACCGGGACATCTACCCAT 937  
Db 1046 AGGCCTTGAGTCATCCCTGGATTGACGGAATACACAGCCCTCCACCGGGACATCTACCCAT 1105  
QY 938 CAGTCA - GCCTCAGATCAGAGAAGCTTTGCTAGAGCAAGTGGAGCAAGCCCTTCAC 996  
Db 1106 CAGTCAGGGCTCCAGATCCAGAGAAGCTTTGCTAGAGCAAGTGGAGCAAGCCCTTCAC 1165  
QY 997 GCAGCAGCTGTGT - GCACACATGAGGAAGCTACACATGAACCTGCACA - GCCCGGGG 1054  
Db 1166 GCAGCAGCTGTGTAGCACATAGGAAGCTTACATGAACCTGCACACGCCCGGGG 1225  
QY 1055 TCCGCCCCAGAGGTGGAGAACAGSCCCTGAAACTCAAGCCCTCAGAAACCTCTAGACCCA 1114  
Db 1226 TCCGCCCCAGAGGTGGAGAACAGSCCCTGAAACTCAAGCCCTCAGAAACCTCTAGACCCA 1285  
QY 1115 GCTCCCTTGATATCACCATCAGGAGGACCTGTCTGACACACAGTGTAGACATCCCTG 1174  
Db 1286 GCTCCCTTGATATCACCATCAGGAGGACCTGTCTGACACACAGTGTAGACATCCCTG 1345  
QY 1175 CCCTGACCCCAATTA - - - CCCTGACAGATGGCGGGCCCACTGCCCTGTGGCAGGT 1231  
Db 1346 CCCTGACCCCAATTAATACCTTGCCAGCATCGCGCGGGCCCACTGCCCTGTGGCAGGT 1405  
QY 1232 CCCTCAACTGCTGTGTCATAGGCTCCCTCCACATCAGCAGCAGCTGTGGCCCATGCATC 1291  
Db 1406 CCCTCAACTGCTGTGTCATAGGCTCCCTCCACATCAGCAGCAGCTGTGGCCCATGCATC 1465  
QY 1292 AGGGGTCCCTGGCGCGGGCCCTGTGGCTGTGCTGTCC - AGCTGCTGAACATTTGGGAGC 1350  
Db 1466 AGGGGTCCCTGGCGCGGGCCCTGTGGCTGTGCTGTCTCCAAAGCTGCTGAACATTTGGGAGC 1525

QY 1351 AAAGGAAGTCCCTCTACTGCTCTGAGCCACACTCTCTCAAAAGGCCAAACAAACACAG 1410  
Db 1526 AAAGGAAGTCCCTCTACTGCTCTGAGCCACACTCTCTCAAAAGGCCAAACAAACACAG 1585  
QY 1411 AACTTCAAGTCGGAGGTCACTGGTACCAAGTTAAAGCCAGTGGCA - GCTCCCACTGCCGGGC 1469  
Db 1586 AACTTCAAGTCGGAGGTCACTGGTACCGGTTAAAGCCAGTGGCATGCTTCCACTGCCGGC 1645  
QY 1470 AGGCAGACTGGAGTCTGCTCATATTGATTCCTGGAGCCTGTGCTATGTCAC - -TG 1527  
Db 1646 AGGCAGACTGGAGTCTGCTCATATTGATTCCTGGAGCCTGTGCTATGTCACGTG 1705  
QY 1528 CAATTTTCAGAGACATATTCAACTCTCTGCTTCTTCCAAACCTGGTGTCTATCCGGCAG 1587  
Db 1706 AATTTTCAGAGACATATTCAACTCTCTGCTCTTCCAAACCTGGTGTCTATCCGGCAG 1765  
QY 1588 AGGAGGAAGCAGCAGCAAGTGGAGCAGGGCTTAGCAGGAGCAGTTTCT - GGCCAGAAGC 1646  
Db 1766 AGGAGGAAGCAGCAGCAAGTGGAGCAGGGCTTAGCAGGAGCAGTTTCTGGCCAGAAGC 1825  
QY 1647 ACCA - - - - GCCTGCTGCCAGCGGGGACGCCCTCATAGGAGGCCAGGAGGAGGCCCAA 1702  
Db 1826 ACCATGCTGCTTGGCAGCTGGGGCAGCCCGCTCATAGGAGGCCAGGAGGAGGCCCAA 1885  
QY 1703 GGCCTAGAACCTT - GTTGAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAG - - 1759  
Db 1886 GGCCTAGAACCTTGGTTGAAGCTGTGAGGAGGAGCGGTGCCACACAGCTTCCAGTG 1945  
QY 1760 - GTCTCCCTGACCTGCTCTATGCCCCACACCCCTAC - - GTGCCCTGGCTCTGTGCAG 1816  
Db 1946 TCTCCCGTGACCTGGCCTGCTCATATGCCACACCCCTACAGTGGCCGTGCTGTGCAG 2005  
QY 1817 TGT - - AGCTAGATAGCTCTCG - COTGGGTCTGTGCTGTTTGTGCTGTAAGAGCTTAA - -TG 1871  
Db 2006 TGTTACGTTAGATAGCTCTCGCCCTGGGTCTGTGCTGTTTGTGTAAGAGCTTAAATGG 2065  
QY 1872 GGCCTGGCAGCTGTGTACCTTCTCAAGCAAAAGCCATAT - - - GGAGCATCTACCC - - 1925  
Db 2066 GCTGGGCAGGCTGTGTACCTTCTCAAGCAAAAGCCCATATGTGAGCATCTATCCAG 2125  
QY 1926 - AGACTCCCCTCTGSCACACACTCACCTCCACCTCTCAAGCCTCCAAACCTCTTGGCCAGA 1984  
Db 2126 AGATCCCACCTGTGTGCACACTCTCTCTCACATCTCTAAAGTCTCCAACCTCTTGGCCAGA 2185  
QY 1985 TTGGGCTCATTAAT - GTGCTGTGCTGCCCATCTG - CATGAATGACAGGAGCTCCCATG 2042  
Db 2186 TTGGGCTCATTAATGGTGTGCTGCCCATCTGCCCATGAATACCGCCAGCTCCCATG 2245  
QY 2043 GTGCTGTGCTCT - GTGAGCTCTTCAAGTTCTTAATCTTAACTTCCAGGATTAAGTCCCAAGT 2101  
Db 2246 GTGCTGTGCTGTGAGCTCTTCAAGTTCTTAATCTTAACTTCCAGGATTAAGTCCCAAGT 2305  
QY 2102 GCCTGTAGACCCAGCCA - - - GCACACTTCTGGGCCCTTCTCCCTGCTCAATCTAAAAGCA 2158  
Db 2306 GCCTGTAGAACCAAGCAAGCAGCAGCTTCTTGGCCCTTCTCCCTGCTCAATCTAAAAGCA 2365  
QY 2159 GTGCCACACCTCCAAA - GTGGGAATAGAAAGTTTCTCATGATGAG - GGCTGCAAGGAAT 2216  
Db 2366 GTGCCACACCTCCAAAAGGTGGAATAGAAAGTTTCTCATGATGAGGAGTGTCAACGTG 2425  
QY 2217 TCTTATCTGTGCCA - - - - - CATGTCTCTCGTGCACA - - - - CACCCAATGGAGTTA 2262  
Db 2426 AACTTACTTAGTCCATGAGTCCAGCAGTGTCTCGTGCACAGCAGCCCAATGGAGCTTA 2485  
QY 2263 ACCTTGAAGTTGAC - - - - TATTTTAAATGTGTGCCAGAG - - - TTTCAATCTGCTCTG - - 2315  
Db 2486 ACCTTGAAGTTGATATATTTAAGTGTGTGCCAGAGGTTCTAGTGTGCTGCTCTGCT 2545  
QY 2316 TTCCCTTTTCTCTCTTGA - - - - - AGTCCACACACCATTTCTGCTCTCC - - - - 2364  
Db 2546 TCCGCTTTTCTCTCTTGAATGTGCCAGTGTGACATCCATATCTTGTGCTGTTCCCGAGG 2605





DT 22-AUG-2001 (first entry)  
XX Nucleotide sequence of a human kinase polypeptide.  
DE Human; kinase; human disease; human disorder; ss.  
KW Homo sapiens.  
XX WO200142435-A2.  
PN 14-JUN-2001.  
PD 07-DEC-2000; 2000WO-US33240.  
XX 07-DEC-1999; 99US-0169428.  
PF (LEXI-) LEXICON GENETICS INC.  
XX Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;  
PI Abuin A, Sands AT;  
XX WPI; 2001-381667/40.  
XX Novel isolated human kinase polynucleotide that shares structural  
PT similarity with animal kinases including calcium/calmodulin-dependent  
PT protein kinases and serine/threonine protein kinases, useful in  
PT therapeutics -  
XX Disclosure; Page 32; 32pp; English.  
XX The present sequence encodes a kinase polypeptide. The kinase  
CC polynucleotides and polypeptides are useful in therapeutic, diagnostic  
CC and pharmacogenetic applications. They are useful for the detection of  
CC mutant kinases, or inappropriately expressed kinases for the diagnosis  
CC of a disease or disorder. They are useful for screening for drugs (or  
CC high throughput screening of combinatorial libraries) effective in the  
CC treatment of symptomatic or phenotypic manifestations of that disease  
CC or disorder. The polynucleotide sequence is useful as a source of  
CC probes and primers, which can be used to screen libraries, isolate  
CC clones, and prepare cloning and sequencing templates.  
XX Sequence 1671 BP; 453 A; 437 C; 448 G; 333 T; 0 other;  
SQ

Query Match 20.6%; Score 504.6; DB 22; Length 1671;  
Best Local Similarity 71.3%; Pred. No. 4.4e-132;  
Matches 695; Conservative 0; Mismatches 274; Indels 6; Gaps 2;  
QY 66 GGCATGGGTGCGAAGAGAGATGACTGCAGTTCCTGGAGAAACACACACCACCAACAT 125  
DB 163 GGCCATGGGCCGGGGAAGCGGAGAGAGAGTCCCTCCTCGAAAGCAAGCTGAAGACAT 222  
QY 126 CCGGAAACCTTCATTTTATGGAGTGTGGATCAGGAGCTTCTCAGAAAGTTTCTCT 185  
DB 223 CAGAGAGATCTTCAGATTCAAGAGACCTCGAACCAGGGCTTTTCGAAGTGGTTT 282  
QY 186 GGTGAAGCAAGACTGACTGGGAAGCTTTTGTCTGAAGTGCATC---AAGAAGTCACC 242  
DB 283 AGCTGAAGAGAGAGCAACTGGCAAGCTTTTGTGTGAAGTGTATCCCTTAAGAAGCGCT 342  
QY 243 TGCCTTCCGGGACAGCAGCTGGAGATGAGATTGCTGTGTTGMAAAGATCAAGCATGA 302  
DB 343 GAAGGCAAGGAAGCAGCATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 402  
QY 303 AAACATTTGACCCCTGGAGGACATCTATGAGAGCACACCACCTACTACTTGGTCATGCA 362  
DB 403 AATATTTGTTGCCCTGGAAGACATTTATGAAGCCCAATACCTGTACTTGGTCATGCA 462  
QY 363 GCTTTGTTCTGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTCTACACAGAA 422  
DB 463 GCTGGTGTCCGGTGGAGAGCTGTTTTCACCGGATGAGTGGAGAGAGGGGTTTATACAGAA 522  
QY 423 GGATGCCAGTCTGGTGTATCAGCAGCTTTGTCGGCAGTGAATACCTACATGAGAATGG 482

Db 523 GGATGCCAGCACTCTGATCCGCCAAGTCTTTGGACGCCGTACTATATCTCCACAGATGGG 582  
QY 483 CATGTCACACAGACTTAAAGCCCGAAACCTGCTTTTACCTTTACCCCTGAAGAACTC 542  
Db 583 CATGTCACACAGACTTAAAGCCCGAAACCTGCTTTTACCTTTACCCCTGAAGAACTC 642  
QY 543 TAAGATCATGATCAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599  
Db 643 CAAAATAATGATCAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702  
QY 600 CACTGCCCTGTGGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659  
Db 703 CACTGCCCTGTGGAACTCCAGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 762  
QY 660 CAAGGCTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719  
Db 763 CAAGGCTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 822  
QY 720 CCCGTTCTATGAAGAAAGGAGTCTTAAGCTTTTTCAGAAAGATCAAGAGGGCTACTATGA 779  
Db 823 TCCTTTTATGATGAAATGACTCCAAGCTCTTTGAGCAGATCCTCAAGGCGGAATATGA 882  
QY 780 GTTTGAGTCTCCATCTGGGATGACATTTCTGAGTCAGCCAGGACTTTATTTGCCACTT 839  
Db 883 GTTTGACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAGACTTCATTCGGAACCT 942  
QY 840 GCTTGAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGCGCTTGAGTCATCCCTGGAT 899  
Db 943 GATGGAGAAAGCCCGAATAAAGATACACGCTGTGAGCAGCAGCTCGGACCCATGGAT 1002  
QY 900 TGACGAAACACGGCCCTCCACCGGGACATCTACCCATGATGAGCTTCAGATCCAGAA 959  
Db 1003 CGCTGGTGACACAGCCCTCAACAAAAACATCCAGAGTCCGTCAGCGCCAGATCCGGA 1062  
QY 960 GAACCTTGTCTAAGAGAGTGGAGCAGGCTTCAAGCAGCAGCTGTGTGTGACCCACAT 1019  
Db 1063 AAATTTTCCCAAGAGCAATGGAGACAGCATTTATGCCACGCCGCTGCTGTGAGACATAT 1122  
QY 1020 GAGGAAGCTACACAT 1034  
Db 1123 GAGAAACTACACCT 1137

RESULT 8  
AAI58638  
ID AAI58638 standard; cDNA; 1733 BP.  
XX AC AAI58638;  
XX DT 22-OCT-2001 (first entry)  
XX DE Human polynucleotide SEQ ID NO 841.  
XX KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX OS Homo sapiens.  
XX PN WO200153312-A1.  
XX PD 26-JUL-2001.  
XX PF 26-DEC-2000; 2000WO-US34263.  
XX PR 21-JAN-2000; 2000US-0488725.  
XX PR 25-APR-2000; 2000US-0552317.  
XX PR 09-JUL-2000; 2000US-0598042.  
XX PR 19-JUL-2000; 2000US-0620312.  
XX PR 03-AUG-2000; 2000US-0653450.



PF XX 20-APR-2001; 2001WO-US12992.  
PR XX 20-APR-2000; 2000US-199021P.  
PR XX 28-APR-2000; 2000US-200226P.  
PR XX 05-MAY-2000; 2000US-202339P.  
PR XX 11-MAY-2000; 2000US-203505P.  
PR XX 18-MAY-2000; 2000US-205564P.  
PR XX 26-MAY-2000; 2000US-207739P.  
PR XX 01-JUN-2000; 2000US-208795P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;  
PI Bandman O, Lu DAM, Lal P, Burford N, Khan FA, Wadia NK, Yao MG;  
PI Patterson C, Burdill JD, Marcus GA, Zingler KA, Recipon SA, Lu Y;  
PI Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;  
PI Walsh RT, Ramkumar J, Borowsky ML, Au-young J, Hillman JL;  
PI Gururajan R;  
XX WPI; 2001-611740/70.  
DR P-PSDB; AAE11768.  
XX Human kinases and nucleic acids, useful for preventing diagnosing and  
PT treating cancers, inflammation and immune disorders -  
XX Claim 5; Page 151-152; 166pp; English.  
XX The present invention relates to human kinases (PKIN) and the nucleic  
CC acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is  
CC used in the prevention, diagnosis and treatment of diseases cancers,  
CC adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,  
CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,  
CC gout, microbial infections, cardiovascular disease and/or inflammation,  
CC myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial  
CC infarction, cataract, growth and development disorder, seizure disorder,  
CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage  
CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.  
CC PKIN may be used to treat disorders associated with decreased PKIN  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of PKIN by expressing inactive proteins or to  
CC supplement the patients own production of PKIN. PKIN nucleic acids may be  
CC used to produce the PKIN polypeptide, by inserting the nucleic acids into  
CC a host cell and culturing the cell to express the protein. PKIN nucleic  
CC acid and its complementary sequences may also be used as DNA probes in  
CC diagnostic assays to detect and quantitate the presence of similar  
CC nucleic acid sequences in samples and therefore which patients may be in  
CC need of restorative therapy. The present sequence is human PKIN-2 cDNA.  
XX Sequence 1736 BP; 463 A; 456 C; 466 G; 351 T; 0 other;

Query Match 20.6%; Score 503; DB 22; Length 1736;  
Best Local Similarity 71.2%; Pred. No. 1.3e-131;  
Matches 694; Conservative 0; Mismatches 275; Indels 6; Gaps 2;  
Qy 66 GGCAATGGGCGAAGGAAGATGACTGCAGTTCTCTGGAAGAAACAGACCACCAACAT 125  
Db 155 GGCCATGGCCGGGAGAACGCGAGAGCAGCTCTCTCTGGAAGAAACAGCTGAAGACAT 214  
Qy 126 CCGAATAACCTTCATTTTTATGGAAGTGTGGGATCAGAGCTTTCTCAGAAGTTTCTCT 185  
Db 215 CAAGAAGATCTTCGAGTCTAAAGAGACCCCTCGGAACGGGGCCCTTTTCCGAGTGGTTT 274  
Qy 186 GGTGAAGCAAGAGTACTGGGAGCTCTTTGCTCTGAGTGCATC- --AAGAAGTCACC 242  
Db 275 AGCTGAAGAGAGCAACTGGCAAGCTCTTTGCTGTGAAGTGATCCCTAAGAAGGCGCT 334  
Qy 243 TGCCTCCGGACAGCAGCCTGGAGATGAGATTGCTGTGTTGAAAGATCAAGCATGA 302  
Db 335 GAAGGGCAAGAAAGCAGCATGAGATGAGATAGCGCTCTCGAAGAGATTAAAGCATCA 394  
Qy 303 AAACATTTGTGACCTGGAGACATCTATGAGACACACCCACTACTACTGTGTCATGCA 362  
Db 395 AAATATTGTTGCCCTGGAAGACATTTATGAAGAGCCCAAAATCATCTGTACTGTGTCATGCA 454

Qy 363 GCTTGTCTTGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAA 422  
Db 455 GCTGTGTCCGGTGGAGAGCTTTTGACCGGATAGTGGAGAGGGGTTTATACAGAA 514  
Qy 423 GGATCCAGTCTGTGTGATCCAGCAGGTCTTGTGCGCAGTGAATACTACATGAGATGG 482  
Db 515 GGATCCAGCAGCTCTGTGATCCCGCAAGTCTTGGACGCCCTCTACTATCTCCACAGATGG 574  
Qy 483 CATCTCCACAGAGACTTAAAGCCCGGAAACCTGCTTTACCTTACCCTTGAAGAGACTC 542  
Db 575 CATCTCCACAGAGACTTAAAGCCCGGAAATCTCTTGTACTACAGTCAAGATGAGAGTCC 634  
Qy 543 TAAGATCATGATCATGACTTGTGCTCTCAAGATGGAACAGAGATGG- --CATCATGCT 599  
Db 635 CAAATATATGATCATGACTTGTGATTTGTCATAAATGAGGGCAAGAGATGTGATGTC 694  
Qy 600 CACTGCCCTGTGGGACCCAGGCTAGCTGGCTCCAGAACTGTGGCCCAAGAACCTACAG 659  
Db 695 CACTGCCCTGTGGAACTCCAGGCTATGCTGCTCTGAAGTCTCGCCCAAGAACCTTACAG 754  
Qy 660 CAAGCTCTGTGATTCCTGCTCATCGGCTCATCACCTACATATTGCTCTGTGGATACCC 719  
Db 755 CAAAGCCGTTGACTGCTGCTCATCGGAGTATTGCTCTACATCTGCTCTCGGCTACCC 814  
Qy 720 CCGCTTCTATGAAGAAACGGAGTCTTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGA 779  
Db 815 TCCTTTTATGATGAAATGACTCCCAAGCTCTTTGAGCAGATCCTCAAGGCGGAATGA 874  
Qy 780 GTTTCAGTCTCCATCTCTGGGATGACATTTCTGAGTCAGCCCAAGGACTTTATTTGCCACTT 839  
Db 875 GTTTCAGTCTCCATCTCTGGGATGACATCTCGGACTCTGCAAAAGACTTCATTCGGAACCT 934  
Qy 840 GCTTGAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGGCGCTTGTAGTCATCCCTGGAT 899  
Db 935 GATGAGAAGGACCGGAAATAAAGATACACGCTGTGAGCAGCAGCTCGGCCACCATGGAT 994  
Qy 900 TGACGGAACAGCGGCTCCACCGGGAGATCTACCCATCAGTCAGCTCCAGCTCCAGAAC 959  
Db 995 CGCTGGTGACACAGCGGCTTCAACAAAACATCCAGAGTCCCTGACGCGCCAGATCCGGAA 1054  
Qy 960 GAACCTTGTGAAGCAAGTGGAGGCAAGCGCTTCAACGACGACGCTGTGGTGCACCAT 1019  
Db 1055 AAACCTTTCGAAGACAAATGGAGACAGCATTTAATCCACGCGGCTCGTGAGACATAT 1114  
Qy 1020 GAGGAAGTACACAT 1034  
Db 1115 GAGAAACTACACCT 1129  
RESULT 10  
AAI60424  
ID AAI60424 standard; cDNA; 2164 BP.  
AC AAI60424;  
XX 22-OCT-2001 (first entry)  
XX Human polynucleotide SEQ ID NO 4413.  
DE Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-brager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
leukaemia; ss.  
XX Homo sapiens.  
XX WO200153312-A1.  
PN 26-JUL-2001.

PF	26-DEC-2000; 2000WO-US34263.	
XX		
XX	21-JAN-2000; 2000US-0488725.	
PR	25-APR-2000; 2000US-0552317.	
PR	09-JUL-2000; 2000US-0598042.	
PR	19-JUL-2000; 2000US-0620312.	
PR	03-AUG-2000; 2000US-0653450.	
PR	14-SEP-2000; 2000US-0662191.	
PR	19-OCT-2000; 2000US-0693036.	
PR	29-NOV-2000; 2000US-0727344.	
XX	(HYSE-) HYSEQ INC.	
XX	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;	
PI	Zhao QA, Zhou P, Goodrich R, Dimanac RT;	
XX	WPI: 2001-442253/47.	
DR	P-PSDB; AAM41268.	
DR		
XX	Novel nucleic acids and polypeptides, useful for treating disorders	
PT	such as central nervous system injuries -	
PT		
XX	Claim 1; SEQ ID NO 4413; 10078pp; English.	
XX		
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and	
CC	the encoded polypeptides (AAM38642-AAM42213) with nootropic,	
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful	
CC	in gene therapy. A composition containing a polypeptide or polynucleotide	
CC	of the invention may be used to treat diseases of the peripheral nervous	
CC	system, such as peripheral nervous injuries, peripheral neuropathy and	
CC	localised neuropathies and central nervous system diseases, such as	
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic	
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the	
CC	utilisation of the activities such as: Immune system suppression,	
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic	
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,	
CC	assays for receptor activity, arthritis and inflammation, leukaemias and	
CC	C.N.S disorders.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification.	
XX		
XX	Sequence 2164 BP; 557 A; 542 C; 560 G; 505 T; 0 other;	
SQ		
	Query Match 20.6%; Score 503; DB 22; Length 2164;	
	Best Local Similarity 71.2%; Pred. No. 1.4e-131;	
	Matches 694; Conservative 0; Mismatches 275; Indels 6; Gaps 2;	
QY	66 GGCAATGGGTGCGAAAGGAAGATGACTGCAGTTCCTCGAAGAAACAGACCACCAACAT 125	
DB	140 GGCCATGGCCCGGAGACGGCGAGAGACGCTCTCTCGAAGAAAGCAAGCTGAAGACAT 199	
QY	126 CCGGAAACCTTCATTTTATGGAAGTGTGGAGTCAGGAGCTTTCAGAAAGTTTCT 185	
DB	200 CAAGAAGATCTTCAGTTCAAAGAGACCCCTCGGAACCGGGCTTTTCGGAAGTGGTTT 259	
QY	186 GGTGAACCAAGACTGACTGGGAGCTCTTCTCTGAGTGCATC---AAGAAGTCACC 242	
DB	260 AGCTGAGAGAAAGGCACTGGCAGCTCTTTGCTGGAAGTGTATCCCTAAGAAGCGGT 319	
QY	243 TGCTTCCTCGGACAGACGCTGGAGATGAGATTGCTGTGTGAAAGATCAAGCATGA 302	
DB	320 GAAGGCAAGGAAGCAGCATAGATGAGATGAGTAGCGCTCTGAGAAAGATTGAAGCATGA 379	
QY	303 AACATTGTGACCTCGAGGACATCTATGAGAGCACCCACCTACTACTGGTATGCA 362	
DB	380 AAATAATTGTGCTCGCTGGAAGACATTTATGAAAGCCCAATACACCTGTACTTGGTATGCA 439	
QY	363 GCTGTGTTCTGTTGGGAGCTCTTTGACCGGATCCTTGGAGCGGGGTGTCTACACAGAA 422	
DB	440 GCTGGTTCGGGTGGAGAGCTGTTTGACCGGATGAGTGGAGAGGGGGTTTTATACAGAGAA 499	
QY	423 GGATGCCAGTCTGGTGATCCAGCAGGTCTTGTGCGCAGTGAATAACCTACATGAGAAATGG 482	

Db

500 GGATGCCAGCACTCTGATCCGCCAAGCTTTGGAGCGCGTGTACTATCTCCACAGAATGGG 559

Qy

483 CATCGTCCACAGAGACTTTAAAGCCCGGAAACCTTGCTTACCTTACCCCTGAAGAGAACTC 542

Db

560 CATCGTCCACAGAGACCTCAAGCCCGGAAATCTCTTGTACTACAGTCAAGATGAGGAGTC 619

Qy

543 TAAGATCATGATCACTGACTTTTGGTCTGTCAGAGATGGAACAGAATGG---CATCATGTC 599

Db

620 CAAATATGATGATGACTGCTTTGGATTTGTCANAANTGAGGGGAAAGAGATGTGATGTC 679

Qy

600 CACTGCCTGTGGGACCCAGGCTACGTGGTCTCCAGAAAGTGTGSCCCAGAAACCTTACAG 659

Db

680 CACTGCCTGTGGAACTCCAGGCTATGTGCTCTCTGAAAGTCTCTGCCCCAGAAACCTTACAG 739

Qy

660 CAAGCGTGTGGATTGCTGGTCCATCGGCGTCACTACATATATTGCTCTGTGGATACCC 719

Db

740 CAAAGCGGTTGACTGCTGGTCCATCGGAGTGATTTGCCCTACATCTTGTCTCGGCTACCC 799

Qy

720 CCCGTTCTATCAAGAAACGGAGTCTAAGCTTTTTCGAGAAGATCAAGGAGGGCTACTATGA 779

Db

800 TCCITTTTATGATGAAATGACTCCAAAGCTCTTTGACGAGATCTCTCAAGCGGGAATATGA 859

Qy

780 GTTTGAGTCTCCATTCTGGGATGACATTTCTGAGTCAGCCCAAGGACTTTTATTTGCCACTT 839

Db

860 GTTTGACTCTCCCTACTGGGATGACATCTCGGACTCTGCAAAAGACTTCATTCGGAACCT 919

Qy

840 GCTTGAGAAGATCCGACGAGCGGTACACTGTGTGAGAAGGCTTTGAGTCATCCCTGGAT 899

Db

920 GATGAGAAGACCGCAATAAAGATACAGCTGTGTGAGCAGCAGCTCGGACCCCATGGAT 979

Qy

900 TGACGGAAACACGCGCTCCACCGGGACATCTACCCATCAGTCAGCTCCAGATCCAGAA 959

Db

980 CGCTGGTGACACAGCCCTCAACAAAACATCCAGAGTCCGTCAGCCCGCAGATCCGGAA 1039

Qy

960 GAACCTTTGCTAAGAGCAAGTGGAGGCAAGCCTTCAACGCGACGAGCTGTGGTGACACCAT 1019

Db

1040 AAACCTTTGCCAAGAGCAAAATGGAGACAAGCATTTAATGCCACGCGCTCTGTCAGACATAT 1099

Qy

1020 GAGGAAGCTACACAT 1034

Db

1100 GAGAAACTACACCT 1114

RESULT 11

AAS06708

ID AAS06708 standard; cDNA; 1074 BP.

XX

AC AAS06708;

XX

DN 12-SEP-2001 (first entry)

XX

DE Polynucleotide sequence encoding human protein kinase #8.

XX

KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;

KW metabolic disorder; immune related disease; neurological disorder;

KW neurodegenerative disorder; inflammatory disorder; infectious disease;

KW reproductive disorder; gene therapy; ss.

XX

OS Homo sapiens.

XX

PN WO200138503-A2.

XX

PD 31-MAY-2001.

XX

PF 22-NOV-2000; 2000WO-US32085.

XX

PR 24-NOV-1999; 99US-0167482.

XX

PA (SUGE-) SUGEN INC.

XX

PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;

PI Flanagan P, Clary D;

XX WPI: 2001-343950/36.  
DR P-PSDB; AAU03508.  
XX Nucleic acids encoding human kinase polypeptides, useful for preventing  
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
PT neuronal-associated diseases, and microbial infections -  
XX  
PS Example 1; Figure 1; 433pp; English.  
XX  
CC AAS06701-1AS06757 encode for novel human protein kinases #1-57. The  
CC novel protein kinases have been identified as members of the tyrosine  
CC or serine/threonine kinase (PK and STK) families. The polynucleotides  
CC encoding protein kinases and the polypeptides may be used in the  
CC prevention, diagnosis and treatment of diseases associated with  
CC inappropriate kinase expression. For example, they may be used to treat  
CC cancers (especially cancers of haematopoietic origin), cardiovascular  
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),  
CC immune related diseases (e.g. rheumatoid arthritis), neurological  
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.  
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious  
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
CC Additionally, polynucleotides encoding protein kinases may be  
CC used for gene therapy and as DNA probes in diagnostic assays.  
CC The protein kinase polypeptides may be used as antigens in the production  
CC of antibodies against the protein kinases and in assays to identify  
CC modulators of protein kinase expression and activity.  
XX  
SQ Sequence 1074 BP; 313 A; 258 C; 276 G; 227 T; 0 other;

Query Match 20.5%; Score 502.2; DB 22; Length 1074;  
Best Local Similarity 71.3%; Pred. No. 1.7e-131;  
Matches 692; Conservative 0; Mismatches 273; Indels 6; Gaps 2;

QY 70 ATGGGTGCGAAGAGAGATGACATGCGAGTCTCTCGAAGAACAGACACCACCAATCCGG 129  
DB 1 ATGGCCGGGAGAGCGGAGAGAGAGTCTCTCGGAAAAAGCAAGCTGAAGACATCAAG 60  
QY 130 AAAACCTTCATTTTATGGAAGTGTGGGATCAGAGAGCTTTCTCAGAAAGTTTCCGTGGT 189  
DB 61 AAGATCTTCGAGTTCAAAGAGAGCCCTCGGAACCGGGGCTTTTCGGAAGTGGTTTAGCT 120  
QY 190 AAGCAAGACTGACTGGGAGCTCTTTGCTCTGAAGTGCATC---AAGAAGTCACCTGCC 246  
DB 121 GAAGAGAGGCAACTGGCAAGCTCTTTGCTGTGAAGTGTATCCCTGAAGAAGCGCGTGAAG 180  
QY 247 TTCCGGGACAGCAGCTGGAGAATGAGATGCTGTGTTGAAAGAGATCAAGCATGAAAC 306  
DB 181 GCGAAGGAAGCAGCATAGAGATGAGATAGCCGTCCTGAGNAAGATTAAAGCATGAAAT 240  
QY 307 ATTGTGACCTGGAGACATCTATGAGAGACACCCACTACTACCTGTGTCATGCGAGTT 366  
DB 241 ATTGTTGCCCTGGAAGACATTTATGAAGCCCAATACCTGTACTTGTGTCATGCGAGCTG 300  
QY 367 GTTCTGTGGGGAGCTCTTTACCGGATCTCGAGCGGGTGTCTACACAGAGAGAT 426  
DB 301 GTGTCGGTGGAGAGCTGTTTACCGCGATAGTGAGAGAGGGGTTTATACAGAGAAGAT 360  
QY 427 GCGAGTCTGGTATCCAGCAGCTCTTTGCGGAGTGAATACCTACATGAGAATGGCATC 486  
DB 361 GCGAGCATCTGATCGCGCAAGCTCTGGACGCGGTGTACTATCTCCACAGATGGGCATC 420  
QY 487 GTCCACAGAGACTTAAAGCCCGAAACCTGCTTTTACCTTACCCCTGAAGAGAACTCTAAG 546  
DB 421 GTCCACAGAGACTCAAGCCCGCAAAATCTCTTGTACTACAGTCAAGATGAGGAGTCCAAA 480  
QY 547 ATCATGATCAGTACTTTGGTGTCTCGAAGATGGAGACAGATGG---CATCATGTCCACT 603  
DB 481 ATAATGATCAGTACTTTGGATTGTCAAAAATGGAGGCAAGAGAGATGTGATGTCCACT 540  
QY 604 GCTGTGGGACCCAGGCTACGTGCTCAGAAAGTGTGGCCAGAAACCCCTACAGCAAG 663  
DB 541 GCTGTGGAACTCCAGGCTATGCTGCTCTCTGAAGTCTTCGCCCCAGAAACCTTACAGCAA 600

QY 664 GCTGTGATTGCTGGTCCATCGCGGTGCATCACCTACATATTGCTCTGTGATACCCCGCG 723  
DB 501 GCCGTTGACTGCTGGTCCATCGGAGTGATGCGCTACATCTTGTCTCGCGCTACCCCTCT 660  
QY 724 TTCTATGAAGAAACGGAGTCTAAAGCTTTTCGAGAAGATCAAGGAGGCGCTACTATGATGTTT 783  
DB 661 TTTTATGATGAAATGACTCTCAAGCTCTTTGAGCAGATCTCAAGGCGGGAATATGAGTTT 720  
QY 784 GAGTCTCCATCTGGGATGACATTTCTGAGTCAGCCCAAGGACTTTATTTGCCACTTCTT 843  
DB 721 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAGACTTCATTCGGAACCTGTATG 780  
QY 844 GAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGGCTTGTAGTCTATCCCTCGATTGAC 903  
DB 781 GAGAAGGACCCGAATAAAGATACACCTGTGAGCAGCGAGCTCGGCACCCATGATCGCT 840  
QY 904 GGAACACGGCCCTCCACCGGACATCTACCCATGATCAGCTCCAGATCCAGATCCAGAGAAC 963  
DB 841 GGTGACACAGCCCTCAACAAAAACATCCAGGAGTCCGTCAGCGCCAGATCCGGAAGAAC 900  
QY 964 TTTGCTAAGAGAGTGGAGGAGCCCTTCAACGACGACGCTGTGTGTGCACCATGAGG 1023  
DB 901 TTTCCCAAGAGCAATGAGACAAAGCAATTTAATGCCACGCGCTCGTGAGACATATGAGA 960  
QY 1024 AAGCTACACAT 1034  
DB 961 AAACCTACACCT 971

RESULT 12  
AAH25119  
ID AAH25119 standard; DNA; 1074 BP.  
XX  
AC AAH25119;  
DF  
DT 22-AUG-2001 (first entry)  
XX  
XX Nucleotide sequence of a human kinase polypeptide.  
XX Human; kinase; human disease; human disorder; ss.  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
CDS 1..1074  
FT /\*tag= a  
FT /transl\_except= "(pos: 838..843, aa: Ala)"  
FT /product= "kinase"  
XX  
PN WO200142435-A2.  
XX  
PD 14-JUN-2001.  
XX  
PF 07-DEC-2000; 2000WO-US33240.  
XX  
PR 07-DEC-1999; 99US-0169428.  
XX  
PA (LEXI-) LEXICON GENETICS INC.  
XX  
XX Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;  
PI Abuin A, Sands AT;  
DR WPI: 2001-381667/40.  
DR P-PSDB; AAB84360.  
XX  
PT Novel isolated human kinase polynucleotide that shares structural  
PT similarity with animal kinases including calcium/calmodulin-dependent  
PT protein kinases and serine/threonine protein kinases, useful in  
PT therapeutics -  
XX  
PS Disclosure; Page 30-31; 32pp; English.  
XX

CC The present sequence encodes a kinase polypeptide. The kinase  
CC polynucleotides and polypeptides are useful in therapeutic, diagnostic  
CC and pharmacogenic applications. They are useful for the detection of  
CC mutant kinases, or inappropriately expressed kinases for the diagnosis  
CC of a disease or disorder. They are useful for screening for drugs (or  
CC high throughput screening of combinatorial libraries) effective in the  
CC treatment of symptomatic or phenotypic manifestations of that disease  
CC or disorder. The polynucleotide sequence is useful as a source of  
CC probes and primers, which can be used to screen libraries, isolate  
CC clones, and prepare cloning and sequencing templates.  
XX  
SQ Sequence 1074 BP; 313 A; 258 C; 276 G; 227 T; 0 other;

Query Match 20.5%; Score 502.2; DB 22; Length 1074;  
Best Local Similarity 71.3%; Pred. No. 1.7e-131;  
Matches 692; Conservative 0; Mismatches 273; Indels 6; Gaps 2;

QY 70 ATGGTCTGAAGGAAGAGTACGCTGCTGCTGGAAGAAACAGACCAACATCCGG 129  
Db 1 ATGGCCGGGAGAACCGGAGAGAGCTCTCTCGAAGAAACAGCTGAAGACATCAAG 60  
QY 130 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAAGTTTCTCGTG 189  
Db 61 AAGATCTTCAGTTCAAAGAGACCTCGGAACCGGGCCCTTTCCGAAAGTGGTTTAGCT 120  
QY 190 AAGCAAGACTGACTGGAGCTCTTTGCTCTGAAGTGCATC--AAGAACTCACTGCC 246  
Db 121 GAAGAGAAGCAACTGGCAAGCTCTTTGCTGTAAGATGATCCCTAAGAGCGCGTGAAG 180  
QY 247 TTCGGGACAGCAGCCTGGAGATGAGATGTCTGTGTAAGAAAGATCAAGCATGAAAC 306  
Db 181 GCGAAGAAAGCAGCATAGAGATGAGATAGCGCTCTGAGAAAGATTAAAGCATGAAAT 240  
QY 307 ATTGTGACCTGGAGACATCTATGAGAGCACACCCACTACTGCTGTCATGCAAGTT 366  
Db 241 ATTGTTGCCCTGGAACATATTATGAAGCCCAAAATCACTGTACTTGTGTCATGCACTG 300  
QY 367 GTTCTGTGGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAGGAT 426  
Db 301 GTGTCGGTGGAGAGCTGTTGACCGGATAGTGGAGAGGGGTTTTATACAGAGAAGAT 360  
QY 427 GCCAGCTGTGGTATCAGCAGGTCTTGTGCGCAGTGAATACCTACATGAGAATGGCATC 486  
Db 361 GCCAGCACTGTGATCGGCCAAGTCTTGGAGCGCGTGTACTATCTCCACAGATGGCATC 420  
QY 487 GTCCAGAGACTTAAGCCCGAAGACCTGCTTTACCTTACCCCTGAAGAGAACTCTAAG 546  
Db 421 GTCCAGAGACCTCAAGCCCGAAGATCTTTGTACTACAGTCAAGATGAGGAGTCCAAA 480  
QY 547 ATCATGACTGACTGCTTGGTCTGTCCAAGATGGAACAAATGG--CATCATGTCCACT 603  
Db 481 ATATGATCAGTACTTTGGATTGTCAAAAATGGAGGGCAAGAGAGATGTGATGCCACT 540  
QY 604 GCCTGTGGACCCAGGCTACGTGGCTCCAGAAAGTCTGGCCGACCAACCCCTACAGCAAG 663  
Db 541 GCCTGTGGAACCTCAGGCTATGTCGCTCTGAAAGTCTCGCCGACCAACCTTACAGCAA 600  
QY 664 GCTGTGATTTGCTGTCATCGCGCTCATCACTACATATGCTCTGTGGATACCCCGG 723  
Db 601 GCGTTGACTGCTGTGTCATCGAGTGAATGTCCTACATCTTCTCTGGGGTACCCTCCT 660  
QY 724 TTCTATGAAGAACCGAGTCTAAGCTTTTCGAAAGATCAAGAGGGGTACTATGAGTTT 783  
Db 661 TTTTATGATGAAATGACTCCAAAGCTCTTTGAGCAGATCTCTCAAGCGGAATATGAGTT 720  
QY 784 GAGTCTCCATCTGGATGACATTTCTGAGTCAGCCAGGACTTTATTTCGCACTGTGTT 843  
Db 721 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAAGACTTCATTTCGGAACCTGATG 780  
QY 844 GAGAAGGATCCGAACAGCGGTACACCTGTGAGAGGGCTTTCAGTCACTCCCTGGATTGAC 903  
Db 781 GAGAAGGCCCGAATAAAGATACAGTGTGAGAGGAGGCTCTCGGCACCCCATGGATCGCT 840

QY 904 GGAAACACGGCCCTCCACCGGACATCTACCCATCAGTCCAGATCCAGATCCAGAAGAC 963  
Db 841 GGTGACACAGCCCTCAACAAAAACATCCAGAGTCCGTGACGCCAGATCCGGAAAAAC 900  
QY 964 TTTCTAAGAGCAAGTGGAGCAAGCCCTTCAAGCGCAGCAGCTGTGTTGCCACCATCAGG 1023  
Db 901 TTTCCCAAGCAATGGAGCAAGCATTTAATGCCACGCCCTCGTGTGACACATATGAGA 960  
QY 1024 AAGTACACAT 1034  
Db 961 AAACATACCT 971

RESULT 13  
AAH25118  
ID AAH25118 standard; DNA; 1158 BP.

XX AAH25118;  
XX 22-AUG-2001 (first entry)  
XX Nucleotide sequence of a human kinase polypeptide.  
XX Human; kinase; human disease; human disorder; ss.  
XX Homo sapiens.

XX Key Location/Qualifiers  
XX CDS 1..1158  
XX FT /\*tag= a  
XX FT /product= "kinase"

XX WO200142435-A2.  
XX 14-JUN-2001.  
XX 07-DEC-2000; 2000WO-US33240.  
XX 07-DEC-1999; 99US-0169428.  
XX (LEXI-) LEXICON GENETICS INC.

XX Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;  
PI Abuin A, Sands AT;  
XX WPI; 2001-381667/40.  
XX P-PSDB; AAB84359.

XX Novel isolated human kinase polynucleotide that shares structural  
XX similarity with animal kinases including calcium/calmodulin-dependent  
XX protein kinases and serine/threonine protein kinases, useful in  
XX therapeutics -  
XX Claim 1; Page 29; 32pp; English.

XX The present sequence encodes a kinase polypeptide. The kinase  
XX polynucleotides and polypeptides are useful in therapeutic, diagnostic  
XX and pharmacogenic applications. They are useful for the detection of  
XX mutant kinases, or inappropriately expressed kinases for the diagnosis  
XX of a disease or disorder. They are useful for screening for drugs (or  
XX high throughput screening of combinatorial libraries) effective in the  
XX treatment of symptomatic or phenotypic manifestations of that disease  
XX or disorder. The polynucleotide sequence is useful as a source of  
XX probes and primers, which can be used to screen libraries, isolate  
XX clones, and prepare cloning and sequencing templates.

XX SQ Sequence 1158 BP; 323 A; 283 C; 304 G; 248 T; 0 other;

Query Match 20.5%; Score 502.2; DB 22; Length 1158;  
Best Local Similarity 71.3%; Pred. No. 1.7e-131;  
Matches 692; Conservative 0; Mismatches 273; Indels 6; Gaps 2;

QY 70 ATGGTCTGAAGGAAGAGTACGCTGCTGGAAGAAACAGACCAACATCCGG 129



[illegible][illegible]

Db 600 GATGCCAGCAGCTCTCATCGCCAGGCTCTGGATGCCGTACTATCTCCACAGAAATGGGC 659  
QY 484 ATCGTCCACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCCTTGAAGAGAACTCT 543  
Db 660 ATTGTCACAGGACCTCAAGCCGAGAACTCTTTACTACAGTCAAGACGAGAGTCC 719  
QY 544 AAGATCATGATCAGTACTGTTGGTCTGTCCAGATGGAACAGAAATGG---CATCATGTCC 600  
Db 720 AAAATAATGATCAGTACTGTTGGTCTGTCCAGAAATGGAGGGCAAGAGAGATGTGATGTC 779  
QY 601 ACTGCTGTGGGACCCAGGCTACCTGAGTCCAGAAAGTCTGGCCCAAGAAACCTTACAGC 660  
Db 780 ACGGCTCGGGGACCCAGGCTATGTGTCGGAAAGTCTCGCCCAAGAAACCTTACAGC 839  
QY 661 AAGGCTGTGGATGCTGTGCTCCATCGGCGTCACTACCTATATGCTCTGTGGATACCCC 720  
Db 840 AAGCTGTGGACTGTGCTCCATCGGGGTGATCGCCTATATCTGCTCTGTGGTTACCCCT 899  
QY 721 CCGTTCATGAAGAACGGAGTCTAAGCTTTTCGAGAAAGATCAAGGAGGGCTACTATGAG 780  
Db 900 CCTTTTATGATGAATGACTCGAAGCTGTTTGAACAGATCCTCAAGGCAGAAATATGAG 959  
QY 781 TTTGACTCTCCATTCTGGGATGACATTTCTGAGTCAGCCAAAGGACTTTATTTGCCACTTG 840  
Db 960 TTTGATTTCCCTACTGGGATGACATCTCCGACTCTGCCAAAGACTTCATTTCGGAATCTG 1019  
QY 841 CTTGAGAAGGATCCGAACGAGCGTACACCTGTGAGAAGGCTTGTAGTCACTCCCTGGATT 900  
Db 1020 ATGGAGAAGACCCCAATAAAGATACACTTGTGAGCAGGAGCTCGACACCCATGGATT 1079  
QY 901 GACGAAACACGGCCCTCCACCGGACATCTACCCAGTCAAGCTCCAGATCCAGAAG 960  
Db 1080 GCTGTGACACAGCCCTTAGCAAAACATTCAGAACTCTGTCAAGTCCCGCAGATCCGGAAG 1139  
QY 961 AACTTTGCTAAGCAAGTGGAGGCAAGCTTCAACGAGCAGCTGTGTCGACCAATG 1020  
Db 1140 AATTTTGAAGAGCAAAATGGAGACAAGCGTTTAAACGCCAGGCAGTCTGTGAGACATATG 1199  
QY 1021 AGGAAGCTACA 1031  
Db 1200 CGGAGGCTCCA 1210

Search completed: March 14, 2003, 14:17:14  
Job time : 574 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 11:30:44 ; Search time 3881 Seconds  
(without alignments)  
10211.378 Million cell updates/sec

Title: US-09-960-643-1  
Perfect score: 2447  
Sequence: 1 tggagtggagctcaagcag.....ttttctctaaaaa 2447

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	Length	ID	Description
1	963.8	39.4	1126	13	BM547443	BM547443 AGENCOURT
2	826	33.8	966	13	BI084897	BI084897 602869466
3	801.4	32.8	1048	13	BI084101	BI084101 602869466
4	794.2	32.5	809	13	BI821474	BI821474 603038366
5	736.8	30.1	740	13	BI824483	BI824483 603038855
6	720.8	29.5	742	13	BI818261	BI818261 603032510

7	720.4	29.4	1068	14	BM921532	BM921532 AGENCOURT
8	713.2	29.1	812	13	BI772626	BI772626 603060879
9	710.2	29.0	1129	14	BM807335	BM807335 AGENCOURT
10	701.4	28.7	824	12	BG715920	BG715920 602676667
11	699	28.6	731	13	BI753035	BI753035 603025844
12	653.6	26.7	905	14	BQ934044	BQ934044 AGENCOURT
13	590.6	24.1	818	12	BG293660	BG293660 602390529
14	555	22.7	656	13	BI667965	BI667965 603282877
15	522	21.3	554	13	BI834635	BI834635 603090418
16	506.8	20.7	573	14	BQ086330	BQ086330 1121607.Y
17	499.4	20.4	501	9	AL134342	AL134342 DKF2P5470
c 18	479	19.6	479	14	BQ102407	BQ102407 1121607.X
19	476.2	19.5	1051	9	AL560091	AL560091 AL560091
c 20	473.2	19.3	484	10	AW016039	AW016039 UI-H-BI0p
c 21	454.2	18.6	555	10	AW251224	AW251224 UI-R-BJ0-
c 22	441.4	18.0	620	14	BM944418	BM944418 UI-M-EH0p
c 23	410	16.8	410	9	AA838372	AA838372 of29a11.s
c 24	405.8	16.6	962	9	AL556476	AL556476 AL556476
25	397.2	16.2	918	14	BQ949629	BQ949629 AGENCOURT
26	378.2	15.5	873	12	BG186668	BG186668 602319906
c 27	371	15.2	451	10	AW254051	AW254051 UI-R-BJ0-
28	360.4	14.7	929	13	BI758466	BI758466 603022650
c 29	358.6	14.7	439	10	AW522224	AW522224 UI-R-B00-
c 30	353	14.4	871	10	BE569540	BE569540 601328368
31	347.4	14.2	1068	14	BM919261	BM919261 AGENCOURT
32	346.2	14.1	627	10	AW826802	AW826802 FK53602.Y
33	345.8	14.1	493	10	AW520558	AW520558 UI-R-BJ0p
34	330.8	13.5	798	12	BG702279	BG702279 602683496
c 35	328.8	13.4	941	9	AL578723	AL578723 AL578723
36	327	13.4	822	13	BI545545	BI545545 603187723
c 37	326.8	13.4	978	9	AL574819	AL574819 AL574819
38	325.2	13.3	556	12	BG895406	BG895406 358836 MA
39	324.6	13.3	1052	14	BQ066920	BQ066920 AGENCOURT
40	321.8	13.2	659	13	BI463145	BI463145 603207913
41	321.8	13.2	758	13	BI761267	BI761267 603044016
42	321.6	13.1	636	10	BE266955	BE266955 601191504
43	321.2	13.1	645	14	BM728430	BM728430 UI-E-EJ0-
44	319.2	13.0	580	9	AJ397261	AJ397261 AJ397261
45	318.8	13.0	748	12	BF143967	BF143967 601791217

ALIGNMENTS

RESULT 1	BM547443	BM547443	1126 bp	mrna	linear	EST 20-FEB-2002
LOCUS	AGENCOURT	6507688	NIH_MGC_125	Homo sapiens	cdna clone	IMAGE:5724450
DEFINITION	5', mRNA sequence.					
ACCESSION	BM547443					
VERSION	BM547443.1	GI:18781213				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .					
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)					
JOURNAL	Unpublished (1999)					
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgaps-r@mail.nih.gov">cgaps-r@mail.nih.gov</a> Tissue Procurement: Invitrogen CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: L1AM12713 Row: 1 Column: 19 High quality sequence stop: 623. Location/Qualifiers 1. .1126					
FEATURES	source					

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5724450"  
/clone\_lib="NIH\_MGC\_125"  
/lab\_host="DH10B"

/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;  
Site\_1: EcoRV (destroyed); Site\_2: NotI; RNA source pool  
of three ovaries, from females ranging in age from 38 to  
49 yo. Library is oligo-dr primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 2.1 kb, insert size range 1-3.5 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 036."

BASE COUNT 299 a 291 c 286 g 244 t 6 others  
ORIGIN

Query Match 39.4%; Score 963.8; DB 13; Length 1126;  
Best Local Similarity 94.3%; Pred. No. 3.le-246;  
Matches 1063; Conservative 0; Mismatches 53; Indels 11; Gaps 6;

QY 2 GGAGTGGGAGCTCAACGACGATTCTTCGGAGTCCCTGSCATCCTCAGAAGCTTCAACTC 61  
DB 1 GGAGTGGGAGCTCAACGACGATTCTTCGGAGTCCCTGSCATCCTCAGAAGCTTCAACTC 60  
QY 62 TGGAGGCAATGGTTCGAAAGGAAGATGACTGCAGTTCCTCGAAGAAACAGACCA 121  
DB 61 TGGAGGCAATGGTTCGAAAGGAAGATGACTGCAGTTCCTCGAAGAAACAGACCA 120  
QY 122 ACATCGGAAACCTTCATTTTATGGAAGTCTGGGATCGGAGCTTCTCAGAAGTTT 181  
DB 121 ACATCGGAAACCTTCATTTTATGGAAGTCTGGGATCGGAGCTTCTCAGAAGTTT 180  
QY 182 TCTGTGTGAAGCAAGACTGACTGGGAAGCTCTTGTCTCTGAAGTGCATCAAGAAGTCAC 241  
DB 181 TCTGTGTGAAGCAAGACTGACTGGGAAGCTCTTGTCTCTGAAGTGCATCAAGAAGTCAC 240  
QY 242 CTGCTTCGGGACAGCAGCCCTGGGAATGAGATTGCTGTGTGTAAGAAATCAAGCATG 301  
DB 241 CTGCTTCGGGACAGCAGCCCTGGGAATGAGATTGCTGTGTGTAAGAAATCAAGCATG 300  
QY 302 ARAACATTGTGACCTGGAGGACATCTATGAGAGCACCACCCACTACTACCTGGTCATGC 361  
DB 301 ARAACATTGTGACCTGGAGGACATCTATGAGAGCACCACCCACTACTACCTGGTCATGC 360  
QY 362 AGCTTTCTGTGTGGGAGCTCTTTGACCGGATCTTGAGGGGGGTGTCTACACAGAGA 421  
DB 361 AGCTTTCTGTGTGGGAGCTCTTTGACCGGATCTTGAGGGGGGTGTCTACACAGAGA 420  
QY 422 AGGATGCCAGTCTGGTGATCCAGCAGGCTTTGTCGGCAGTGAATACCTACATGAGAATG 481  
DB 421 AGGATGCCAGTCTGGTGATCCAGCAGGCTTTGTCGGCAGTGAATACCTACATGAGAATG 480  
QY 482 GCATCTGCACAGAGACTTAAGCCCGAAGAACTCTTTACCTTACCTTACCTTACCTTACCT 541  
DB 481 GCATCTGCACAGAGACTTAAGCCCGAAGAACTCTTTACCTTACCTTACCTTACCTTACCT 540  
QY 542 CTAAGATCATGATCACTGACTTTGGTCTGTCCAAAGATGGAAGAGATGGCATCATGTCCA 601  
DB 541 CTAAGATCATGATCACTGACTTTGGTCTGTCCAAAGATGGAAGAGATGGCATCATGTCCA 600  
QY 602 CTGCTGTGGGACCCAGGCTACGTGGCTCCAGAGTGTGCGCCAGAAACCTTACAGCA 661  
DB 601 CTGCTGTGGGACCCAGGCTACGTGGCTCCAGAGTGTGCGCCAGAAACCTTACAGCA 660  
QY 662 AGGCTGTGATGCTGGTCCATCGCGTCATCACCTACATATTGCTCTGTGATACCCCC 721  
DB 661 AGGCTGTGATGCTGGTCCATCGCGTCATCACCTACATATTGCTCTGTGATACCCCC 720  
QY 722 CGTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAGT 781  
DB 721 CATTCTATCAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAGT 780

QY 782 TTGAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGCCAGGAGCTTATTTTGCCAC--TTG 840  
DB 781 TTGAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGCCAGGAGCTTATTTTGCCACTTT 840  
QY 841 CTTGAGGAGGATCCGACGAGCGGTACACCTGT--GAGNAGGCCCTTGAGTCATCCC--TG 896  
DB 841 CTTGAGGAGGATCCGACGAGCGGTACACCTGTGAGNAGGCCCTTGAGTCATCCCCTGG 900  
QY 897 GATTGACGGAACACAGGCGCC--TCCACGGGACATCTACCCATCAGTCAGCTCCAGATC 954  
DB 901 GATTGACGGAACACAGGCGCCCTCCACGGGNACATCTACCCATCAGTCAGCTCCAGATC 960  
QY 955 CAGAAGAACTTTGTTAAGAGCAAGTGGAGCAAGCCTTCAACGAGCAGAGCTGTGGTCCAC 1014  
DB 961 CAGAAGAACTTTGGCTTAANAACCAATGGGAGGACGCTTCAACCCANCAAGCTGTGGTCCC 1020  
QY 1015 CACATGAGGAGCTTACATGAACTGCACAGCCCGGCGCTCCGCC--AGAGGTGGAG 1071  
DB 1021 CCCCTGAGGAAGCTTAC--CCTGAACCTGGACAGCGCGGAATTCNNCCCCAAAAGGGGGA 1079  
QY 1072 AACAGGCGCCTGAAACTCAAGCCTCAGAACTCTAGACCAGCTC 1118  
DB 1080 AACAGGCGCCTGAAATCCAGCCTCAAAACCTTTGAACCCAGCTTC 1126  
RESULT 2  
BI084897/c 966 bp mRNA linear EST 20-JUN-2001  
LOCUS 602869466T1 NIH\_MGC\_102 Homo sapiens cDNA clone IMAGE:5013856 3',  
DEFINITION mRNA sequence.  
ACCESSION BI084897  
VERSION BI084897.1 GI:14503227  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 966)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-re@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: CLONETECH Laboratories, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1CMI820 row: d column: 17  
High quality sequence stop: 832.  
FEATURES  
Location/Qualifiers  
source 1..966  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5013856"  
/clone\_lib="NIH\_MGC\_102"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: salivary gland; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dr priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed  
by Ling Hong in the laboratory of Gerald M. Rubin  
(University of California, Berkeley) using ZAP-cDNA  
synthesis kit (Stratagene) and superscript II RT (Life  
Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT 219 a 238 c 302 g 207 t  
ORIGIN  
Query Match 33.8%; Score 826; DB 13; Length 966;  
Best Local Similarity 95.9%; Pred. No. 2.le-209;

		Matches	913;	Conservative	0;	Mismatches	30;	Indels	9;	Gaps	6;
QY	1440	TAAAGCCAGTGGCAGCTCCCACTGCGCGCAGGCGAGAGTGTCTCTCATTTATGTG 1499									
Db	952										
QY	1500	ATTCTCTGGAGCCTGTGCCTATGTCACTGC--AATTTTCAGGAGACATATTCAACTCCT--CT 1557									
Db	893										
QY	1558	GCTCTTCCAAA---CCTGGTGTCTATPCGGC--AGAGGGAGGAAGCAGAGCAAGTGGAG 1612									
Db	833	GCTCTTCCAAAAGACCTGGGTGTCTATCCGCCAGAGGGAGGAGGAGCAGCAAGTGGAG 774									
QY	1613	CAGGGCTTAGCAGGACGCTTCTGGCCAGAGCACCACCTCTCTGCCAGCGGGCAGCC 1672									
Db	773										
QY	1673	CCTCATAGGAGGCCAGGAGGCCCAAGGCGTAGAAGCCTTGTGAAGCTGTGAGCA 1732									
Db	713										
QY	1733	GGAGAGCGGTGCCACACAGCTTCCAGTCTCCCTGACCTGCTCTATGCCCCACAC 1792									
Db	653										
QY	1793	CCTACGTGGCGTGGCTCTGTGCAGTGTACGTAGTAGCTCTCGCCTGGGTCTGTCTGTT 1852									
Db	593										
QY	1853	TGTCGTGAAGAAGTTAATGGGTGGCCAGGCTGTGTCACTTCTCCAAAGCAAGCCATAT 1912									
Db	533										
QY	1913	GGAGCATCTACCAGACTCCCACTCTGACACACACTCTCCACCTCTCAAGCCTCCAAAC 1972									
Db	473										
QY	1973	CTCTTGGCCAGATTGGCTCATTATGCTGTTGGCTTGGCTGCCATCTGCATGAATGACAGCA 2032									
Db	413										
QY	2033	GCTCCCATGGTGGTCTGTGAGCTCTTCAAGTTCTTAATCTTAACTCCAGGATTAG 2092									
Db	353										
QY	2093	CTCCCAAGTGGCTGAGACCCAGCCAGCACACTCTTGGGCCCTTCTCCCTGCTCAATCTA 2152									
Db	293										
QY	2153	AAAGCAGTGCCACACCCCTCCAAGTGGATAGAAAGATTTCATGAGTAAGGGCTGCAG 2212									
Db	234										
QY	2213	GAATCTTATCTCTGGCCACATGTCTCTCCGTGCACACACCCCAATGAGATTAACTTTGGAAG 2272									
Db	174										
QY	2273	TTGACTATTTTAAATGTCTGCCAGGAGTTCTAATCTGCTCTGTGTTCCCTTTCTCTCCTT 2332									
Db	114										
QY	2333	GAAAGTCCAGCACACCATTTCTTCTCTTCCCGAGTTTCTCTCGCCTCCACCC 2384									
Db	54										

RESULT 3  
BI084101 1048 bp mRNA linear EST 20-JUN-2001  
LOCUS 602869466F1 NTH\_MGC\_102 Homo sapiens cDNA clone IMAGE:5013856 5',  
DEFINITION mRNA sequence.  
ACCESSION BI084101

		BI084101.1	GI:14502431
VERSION	EST.		
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 1048)		
AUTHORS	NTH-MGC http://mgc.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: ccapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: CLONETECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L1CM1820 row: d column: 17 High quality sequence stop: 853.		
FEATURES	Location/Qualifiers		
source	1..1048		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5013856"		
	/clone_lib="NTH_MGC_102"		
	/tissue_type="epidermoid carcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NTH_MGC Library."		
BASE COUNT	231 a 298 c 264 g 254 t		
ORIGIN	1 others		
Query Match	32.8%;	Score	801.4; DB 13; Length 1048;
Best Local Similarity	93.1%;	Pred. No.	8.1e-203;
Matches	885;	Conservative	0; Mismatches 56; Indels 10; Gaps 4;
QY	1425	GGTCATGGTACCAAGCCAGTGGCAGCTCCCACTGCCGGCAGGCGAGACTGGAGT 1484	
Db	2	GGTCATGGTACCAAGCCAGTGGCAGCTCCCACTGCCGGCAGGCGAGACTGGAGT 61	
QY	1485	CTGTCTCATTTATGTATTCCTGGAGCCTGTGCCTATGTCACTGCAATTTTCAGGAGACAT 1544	
Db	62	CTGTCTCATTTATGTATTCCTGGAGCCTGTGCCTATGTCACTGCAATTTTCAGGAGACAT 121	
QY	1545	ATTCAACTCTCTCTCTTCCAAACCTGGTGTCTATCCGGCAGAGGAGGAGGAGGAGCAGC 1604	
Db	122	ATTCAACTCTCTCTCTTCCAAACCTGGTGTCTATCCGGCAGAGGAGGAGGAGGAGCAGC 181	
QY	1605	AAGTGGAGCAGGCTTAGCAGGACAGTTTCTGGCCAGAACACCAGCCTGCTGCCAGCG 1664	
Db	182	AAGTGGAGCAGGCTTAGCAGGACAGTTTCTGGCCAGAACACCAGCCTGCTGCCAGCG 241	
QY	1665	GGGAGAGCCCTCATAGGAGGCCAGGAGGCCAGGAGGCCAGGAGGCCAGGAGGCCAGGAG 1724	
Db	242	GGGAGAGCCCTCATAGGAGGCCAGGAGGCCAGGAGGCCAGGAGGCCAGGAGGCCAGGAG 301	
QY	1725	TGTGAGCAGGAGAGCGGTGCCACCAGCTTCCAGGTCTTCCCTGACCTGCCCTGCTCTATG 1784	
Db	302	TGTGAGCAGGAGAGCGGTGCCACCAGCTTCCAGGTCTTCCCTGACCTGCCCTGCTCTATG 361	
QY	1785	CCCAACACCCCTACGTGCCGTGGCTCTGTGAGTGTACGTAGTAGTCTGCCCTGGGTCT 1844	
Db	362	CCCAACACCCCTACGTGCCGTGGCTCTGTGAGTGTACGTAGTAGTCTGCCCTGGGTCT 421	
QY	1845	GTGCTGTTTGTGCTGAAAAGCTTAATGGGCTGGCCAGGCTGTGTACACCTTCTTCCAGCAA 1904	

|||||  
Db 422 GTGCTGTTGTCTGTAAGAAAGCTTAATGGGCTGGCCAGGCTGTGCACCTTCTCCAAAGCAA 481  
QY 1905 AGCCATATGGAGACTACCCAGAGTCCCACTCTGCACACACTCACTCCCACTCTCAAG 1964  
Db 482 AGCCATATGGAGACTACCCAGAGTCCCACTCTGCACACACTCACTCCCACTCTCAAG 541  
QY 1965 CTTCAACCTCTGGCCAGATTGGGCTCATTAATCTGCTTGGCTGCCCATCTGCATGAAT 2024  
Db 542 CTTCAACCTCTGGCCAGATTGGGCTCATTAATCTGCTTGGCTGCCCATCTGCATGAAT 601  
QY 2025 GACAGGAGCTCCCATGGTGTCTGCTGTGAGCTTCAAGTTCATTAATCTTAACATCC 2084  
Db 602 GACAGGAGCTCCCATGGTGTCTGCTGTGAGCTTCAAGTTCATTAATCTTAACATCC 661  
QY 2085 AGSATTAGCTCCCAAGTGGCTGAGACCCAGCCAGCACACTTCTGGCCCTTCTCCCTGCC 2144  
Db 662 AGSATTAGCTCCCAAGTGGCTGAGACCCAGCCAGCACACTTCTGGCCCTTCTCCCTGCC 718  
QY 2145 TCAATCTAAAGCACTGGCCACACCCCTCCAAAGTGAATAGAAAGTTCATGAGTAA-- 2202  
Db 719 TCAATCTAAAGCACTGGCCACACCCCTCCAAAGTGAATAGAAAGTTCATGAGTAA 778  
QY 2203 -GGGTGCAAGAAATCTTATCTGCGGCACATGCTCCTCGTGCACACACCAATGGAGTT 2261  
Db 779 NGGCTGGCCAGAAATCTTATCTGCGGCACATGCTCCTCGTGCACACACCAATGGAGT 838  
QY 2262 AACCTTGG----AAGTGACTATTTAATGCTGCGCAGGAGTTCTAATCTCGCTCTGTT 2317  
Db 839 TACCCTTGGAAAGTGTACCTATTTAATGCTTTCGCGGAGTCTAATCTGCGCTCG 898  
QY 2318 CCCTTTCTCTCTGGAAGTCCAGCACACCACTTCTGCTTCCCTCCCAAGTT 2368  
Db 899 GTCTTTCTCTCTGGAAGTCCAGCACACCACTTCTGCTTCCCTCCCAAGTT 949

RESULT 4  
BI821474  
LOCUS 603038366F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5179336 5',  
DEFINITION mRNA sequence.  
ACCESSION BI821474.1 GI:15933024  
VERSION EST.  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 809)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LAM11447 row: c column: 17  
High quality sequence stop: 808.  
Location/Qualifiers  
1..809  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5179336"  
/lab\_host="NIH\_MGC\_115"  
/lab\_host="DH10B"  
/note="Organ: pooled brain, lung, testis; Vector:  
pCMV-SPORT6; Site:1; NotI; Site:2; EcoRV (destroyed); RNA  
source anonymous pool of 6 male brains, age range 23-27; 1

male lung, age 27; and 1 male testis, age 69. Library is  
oligo-dT primed and directionally cloned (EcoRV site is  
destroyed upon cloning). Average insert size 1.8 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by C.  
Gruber (Invitrogen). Research Genetics tracking code  
021. Note: this is a NIH\_MGC Library."  
BASE COUNT 212 a 192 c 218 g 187 t  
ORIGIN  
Query Match 32.5%; Score 794.2; DB 13; Length 809;  
Best Local Similarity 99.6%; Pred. No. 5.9e-201;  
Matches 796; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 TGGAGTGGGAGCTCAACGAGGATTTCTCCGAGTCCCTGGCATCCTCAGAAAGCTTCAACT 60  
Db 6 TGGAGTGGGAGCTCAACGAGGATTTCTCCGAGTCCCTGGCATCCTCAGAAAGCTTCAACT 65  
QY 61 CTGAGGCAATGGGTGCAAGGAAGAGATGACTGCAAGTTCTTGGGAAGAACAGACCAACC 120  
Db 66 CTGAGGCAATGGGTGCAAGGAAGAGATGACTGCAAGTTCTTGGGAAGAACAGACCAACC 125  
QY 121 AACATCCGGAACCTTCATTTTATGGAAGTGTGGGATCAGGAGCTTTCTCAGAGTT 180  
Db 126 AACATCCGGAACCTTCATTTTATGGAAGTGTGGGATCAGGAGCTTTCTCAGAGTT 185  
QY 181 TTCTGCTGGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA 240  
Db 186 TTCTGCTGGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA 245  
QY 241 CTGCTCTCGGGACAGCAGCTGGGAAGTGAATGAGATGCTGTGTGAAAAGATCAAGCAT 300  
Db 246 CTGCTCTCGGGACAGCAGCTGGGAAGTGAATGAGATGCTGTGTGAAAAGATCAAGCAT 305  
QY 301 GAAACATTTGTACCCCTGGAGGACATCTATGAGAGACACCCACTACTACCTGGTCAATG 360  
Db 306 GAAACATTTGTACCCCTGGAGGACATCTATGAGAGACACCCACTACTACCTGGTCAATG 365  
QY 361 CAGCTTCTTCTGCTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAG 420  
Db 366 CAGCTTCTTCTGCTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAG 425  
QY 421 AAGGATCCAGTCTGGTGATCCAGCAGGCTTGTCTGGCAGTGAATACCTACATGAGAAT 480  
Db 426 AAGGATCCAGTCTGGTGATCCAGCAGGCTTGTCTGGCAGTGAATACCTACATGAGAAT 485  
QY 481 GGCATCTCCACAGAGACTTTAAAGCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC 540  
Db 486 GGCATCTCCACAGAGACTTTAAAGCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC 545  
QY 541 TCTAAGATCATGATCACTGACTTTGGTCTGTCCAAGATGGAACAGAAATGCGATCATGTCC 600  
Db 546 TCTAAGATCATGATCACTGACTTTGGTCTGTCCAAGATGGAACAGAAATGCGATCATGTCC 605  
QY 601 ACTGCCCTGGGACCCAGCAGCTAGTGGCTCCAGAGTGTGCCAGAACCCCTACAGC 660  
Db 606 ACTGCCCTGGGACCCAGCAGCTAGTGGCTCCAGAGTGTGCCAGAACCCCTACAGC 665  
QY 661 AAGGCTGTGAGTGTGCTGGTCCATCGGCGTGTATGCTTGTCTGTGGATACCCCT 720  
Db 666 AAGGCTGTGAGTGTGCTGGTCCATCGGCGTGTATGCTTGTCTGTGGATACCCCT 725  
QY 721 CCGTCTTATGAAGAACGGAGTCTAAGCTTTTCGAGAAAGATCAAGAGGCGCTACTATGAG 780  
Db 726 CCGTCTTATGAAGAACGGAGTCTAAGCTTTTCGAGAAAGATCAAGAGGCGCTACTATGAG 785  
QY 781 TTTGAGTCTCCATTCTGGG 799  
Db 786 TTGAGTCTCCATTCTGGG 804  
RESULT 5  
BI824483

```
LOCUS      BI824483              740 bp      mRNA      linear      EST 04-OCT-2001
DEFINITION 603038855F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179957 5',
mRNA sequence.
ACCESSION  BI824483
VERSION    BI824483.1  GI:15936033
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 740)
            NIH-MGC http://mgi.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1448 row: m column: 14
            High quality sequence stop: 709.
            Location/Qualifiers
                Location: 740
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5179957"
                /clone_lib="NIH_MGC_115"
                /lab_host="DH10B"
                /note="Organ: pooled brain, lung, testis; Vector:
                pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                source anonymous pool of 6 male brains, age range 23-27; 1
                male lung, age 27; and 1 male testis, age 69. Library is
                oligo-dT primed and directionally cloned (EcoRV site is
                destroyed upon cloning). Average insert size 1.8 kb,
                insert size range 1-3 kb. Library is normalized and
                enriched for full-length clones and was constructed by C.
                Gruber (Invitrogen). Research Genetics tracking code
                021. Note: this is a NIH_MGC Library."

BASE COUNT  196 a 181 c 196 g 167 t

ORIGIN
Query Match      30.1%; Score 736.8; DB 13; Length 740;
Best Local Similarity 99.7%; Pred. No. 1.3e-185;
Matches 738; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  1  TGGAGTGGGAGCTCAAGCAGGATTCCTCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT 60
Db  1  TGGAGTGGGAGCTCAAGCAGGATTCCTCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT 60
QY  61  CTGAGGCAATGGGTCCGAAGGAAGAGATGACTGCAGTTCCTGGGAAGAACAGACCACC 120
Db  61  CTGAGGCAATGGGTCCGAAGGAAGAGATGACTGCAGTTCCTGGGAAGAACAGACCACC 120
QY  121  AACATCGGAACACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTT 180
Db  121  AACATCGGAACACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTT 180
QY  181  TTCTGTGTGAAGCAAAAGACTGACTGGGAAGCTTTTGTGCTCTGAAGTGCATCAAGAAGTCA 240
Db  181  TTCTGTGTGAAGCAAAAGACTGACTGGGAAGCTTTTGTGCTCTGAAGTGCATCAAGAAGTCA 240
QY  241  CTTGCCCTTCGGGACAGCAGCCTGGAGAATGAGATTGCTGTGTTGAAAAGATCAAGCAT 300
Db  241  CTTGCCCTTCGGGACAGCAGCCTGGAGAATGAGATTGCTGTGTTGAAAAGATCAAGCAT 300
QY  301  GAAAACATTGTGACCTTGAGGACATCTATGAGAGCACCACCCTACTACTACCTGGTCATG 360
Db  301  GAAAACATTGTGACCTTGAGGACATCTATGAGAGCACCACCCTACTACTACCTGGTCATG 360

LOCUS      BI818261              742 bp      mRNA      linear      EST 04-OCT-2001
DEFINITION 603032510F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173587 5',
mRNA sequence.
ACCESSION  BI818261
VERSION    BI818261.1  GI:15928724
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 742)
            NIH-MGC http://mgi.nci.nih.gov/
AUTHORS    National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished (1999)
JOURNAL    Contact: Robert Strausberg, Ph.D.
COMMENT    Email: cgapbs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1432 row: d column: 04
            High quality sequence stop: 742.
            Location/Qualifiers
                Location: 742
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5173587"
                /clone_lib="NIH_MGC_115"
                /lab_host="DH10B"
                /note="Organ: pooled brain, lung, testis; Vector:
                pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                source anonymous pool of 6 male brains, age range 23-27; 1
                male lung, age 27; and 1 male testis, age 69. Library is
                oligo-dT primed and directionally cloned (EcoRV site is
                destroyed upon cloning). Average insert size 1.8 kb,
                insert size range 1-3 kb. Library is normalized and
                enriched for full-length clones and was constructed by C.
                Gruber (Invitrogen). Research Genetics tracking code
                021. Note: this is a NIH_MGC Library."

FEATURES             source
LOCUS      BI818261              742 bp      mRNA      linear      EST 04-OCT-2001
DEFINITION 603032510F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173587 5',
mRNA sequence.
ACCESSION  BI818261
VERSION    BI818261.1  GI:15928724
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 742)
            NIH-MGC http://mgi.nci.nih.gov/
AUTHORS    National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished (1999)
JOURNAL    Contact: Robert Strausberg, Ph.D.
COMMENT    Email: cgapbs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1432 row: d column: 04
            High quality sequence stop: 742.
            Location/Qualifiers
                Location: 742
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5173587"
                /clone_lib="NIH_MGC_115"
                /lab_host="DH10B"
                /note="Organ: pooled brain, lung, testis; Vector:
                pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                source anonymous pool of 6 male brains, age range 23-27; 1
                male lung, age 27; and 1 male testis, age 69. Library is
                oligo-dT primed and directionally cloned (EcoRV site is
                destroyed upon cloning). Average insert size 1.8 kb,
                insert size range 1-3 kb. Library is normalized and
                enriched for full-length clones and was constructed by C.
                Gruber (Invitrogen). Research Genetics tracking code
                021. Note: this is a NIH_MGC Library."

```

BASE COUNT ORIGIN	197 a	180 c	197 g	168 t
Query Match	29.5%; Score 720.8; DB 13; Length 742;			
Best Local Similarity	99.6%; Pred. No. 2.4e-181;			
Matches 733; Conservative	0; Mismatches 2; Indels 1; Gaps 1;			
Qy	1	TGGAGTGGGAGTCAA-GCAGGATTCTTCCCGAGTTCCTCGGATCCCTCGCATCCCTCAGAAGTTCACAC	59	
Db	7	TGGAGTGGGAGTCAATCAGGATTCTTCCCGAGTTCCTCGGATCCCTCGCATCCCTCAGAAGTTCACAC	66	
Qy	60	TCTGGAGCAATGGGTGAAAGGAAAGATGACTGCGATTGCTTCCCTGGGAAGAAACAGACAC	119	
Db	67	TCTGGAGCAATGGGTGAAAGGAAAGATGACTGCGATTGCTTCCCTGGGAAGAAACAGACAC	126	
Qy	120	CAACATCCGGAAACCTTTCATTTTATGGAAGTGTCTGGATCAGGAGCTTCTCAGAAGT	179	
Db	127	CAACATCCGGAAACCTTTCATTTTATGGAAGTGTCTGGATCAGGAGCTTCTCAGAAGT	186	
Qy	180	TTTCTGTGTAAGCAAGACGACTGCTGGGAAGTCTTTTCTCTGAAGTGCATCAAGATC	239	
Db	187	TTTCTGTGTAAGCAAGACGACTGCTGGGAAGTCTTTTCTCTGAAGTGCATCAAGATC	246	
Qy	240	ACCTGCCCTCCGGACACAGCCTGGAGAATGAGATTGCTGTGTGAAAGATCAAGCA	299	
Db	247	ACCTGCCCTCCGGACACAGCCTGGAGAATGAGATTGCTGTGTGAAAGATCAAGCA	306	
Qy	300	TGAAACATTTGTGACCTTGGAGGACATCTATGAGAGCACCAACCTACTACTCTGGTGCAT	359	
Db	307	TGAAACATTTGTGACCTTGGAGGACATCTATGAGAGCACCAACCTACTACTCTGGTGCAT	366	
Qy	360	GCAGCTTGTCTGTGGGGAGTCTTTTGACCGGATCCTGGAGCGGGGTGCTACACAGA	419	
Db	367	GCAGCTTGTCTGTGGGGAGTCTTTTGACCGGATCCTGGAGCGGGGTGCTACACAGA	426	
Qy	420	GAAGATCCAGTCTGTGATCCAGCAGTCTTCTCGCAGTGAATACCTACATGAGAA	479	
Db	427	GAAGATCCAGTCTGTGATCCAGCAGTCTTCTCGCAGTGAATACCTACATGAGAA	486	
Qy	480	TGGCATCTGCCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAA	539	
Db	487	TGGCATCTGCCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAA	546	
Qy	540	CTCTAAGATCATGATCACTGCTTGGTCTGTCCAGATGGAACAGATGGCATCATGTC	599	
Db	547	CTCTAAGATCATGATCACTGCTTGGTCTGTCCAGATGGAACAGATGGCATCATGTC	606	
Qy	600	CACTGCCTGTGGACCCAGAGTCTAGTGGCTCCAGAGTGTGGCCAGAAACCTCAG	659	
Db	607	CACTGCCTGTGGACCCAGAGTCTAGTGGCTCCAGAGTGTGGCCAGAAACCTCAG	666	
Qy	660	CAAGGCTGTGGATTGCTGGTCCATCGGCGTCACTACCTACATATTTGCTCTGTGGATACCC	719	
Db	667	CAAGGCTGTGGATTGCTGGTCCATCGGCGTCACTACCTACATATTTGCTCTGTGGATACCC	726	
Qy	720	CCCGTTCTATGAAGAA	735	
Db	727	CCCGTTCTATGAAGAA	742	
RESULT 7				
BM921532				
LOCUS	BM921532 1068 bp mRNA linear EST 12-MAR-2002			
DEFINITION	AGENCOURT_6708041 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5753010			
	5', mRNA sequence.			
ACCESSION	BM921532			
VERSION	BM921532.1 GI:19371911			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1068)			

AUTHORS  
TITLE  
JOURNAL  
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM12788 row: b column: 19  
High quality sequence stop: 592.  
Location/Qualifiers

FEATURES

source

1..1068  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5753010"  
/clone\_lib="NIH\_MGC\_115"  
/lab\_host="DH10B"  
/note="Organ: pooled brain, lung, testis; Vector:  
pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA  
source anonymous pool of 6 male brains, age range 23-27; 1  
male lung, age 27; and 1 male testis, age 69. Library is  
oligo-dr primed and directionally cloned (EcoRV site is  
destroyed upon cloning). Average insert size 1.8 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by C.  
Gruber (Invitrogen). Research Genetics tracking code  
021. Note: this is a NIH\_MGC Library."

BASE COUNT  
ORIGIN

279 a 281 c 260 g 248 t  
Query Match 29.4%; Score 720.4; DB 14; Length 1068;  
Best Local Similarity 98.9%; Pred. No. 3.8e-181;  
Matches 746; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

Qy	1	TGGAGTGGGAGTCAACGAGGATTCTTCCGAGTTCCTCGGATCCCTCGCATCCCTCAGAAGTTCACACT	60	
Db	14	TGGAGTGGGAGTCAACGAGGATTCTTCCGAGTTCCTCGGATCCCTCGCATCCCTCAGAAGTTCACACT	73	
Qy	61	CTGGAGCAATGGTCCGAAAGGAAAGATGACTGCAGTTCCTGGAGAAACAGACACC	120	
Db	74	CTGGAGCAATGGTCCGAAAGGAAAGATGACTGCAGTTCCTGGAGAAACAGACACC	133	
Qy	121	AACATCCGGAACCTTCATTTTATGGAAGTGTCTGGATCAGGAGCTTTCTCAGAAGTT	180	
Db	134	AACATCCGGAACCTTCATTTTATGGAAGTGTCTGGATCAGGAGCTTTCTCAGAAGTT	193	
Qy	181	TTCTGTGTGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA	240	
Db	194	TTCTGTGTGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA	253	
Qy	241	CCTGCCCTCCGGACACGACCTGGGAATGAGATTGCTGTGTTGAAAAGATCAAGCAT	300	
Db	254	CCTGCCCTCCGGACACGACCTGGGAATGAGATTGCTGTGTTGAAAAGATCAAGCAT	313	
Qy	301	GAAGCAATTGTGACCTGGAGGACATCTATGAGAGCACCAACCCACTACTACCTGGTCATG	360	
Db	314	GAAGCAATTGTGACCTGGAGGACATCTATGAGAGCACCAACCCACTACTACCTGGTCATG	373	
Qy	361	CAGCTTGTCTTGTGGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGCTTCACAGAG	420	
Db	374	CAGCTTGTCTTGTGGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGCTTCACAGAG	433	
Qy	421	AAGATGCCAGTCTGGTGATCCAGCAGCTTTGTGCGCAGTGAATACCTACATGAGAAT	480	
Db	434	AAGATGCCAGTCTGGTGATCCAGCAGCTTTGTGCGCAGTGAATACCTACATGAGAAT	493	
Qy	481	GGCATCTGCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC	540	
Db	494	GGCATCTGCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC	553	

```
QY 541 TCTAAGATCATGATGACTTTGGTCTGTCCAGATGGAACAGAAATGGCATCATGTCC 600
Db 554 TCTAAGATCATGATGACTTTGGTCTGTCCAGATGGAACAGAAATGGCATCATGTCC 613
QY 601 A-CTGGCTGTGGACCCAGGCTACCTGGCTCCAGAAAGTGTGGCCAGAAACCCCTACAG 659
Db 614 ACCTGCTGTGGACCCAGGCTACCTGGCTCCAGAAAGTGTGGCCAGAAACCCCTACAG 673
QY 660 CAAGGCTGTGGATTTGCTGTGCTCCATCGGCGTCACTACATATTTGCTGTGGATACCC 719
Db 674 CAAGGCTGTGGATTTGCTGTGCTCCATCGGCGTCACTACATATTTGCTGTGGATACCC 733
QY 720 CCGG-TTCTATGAAGAACGGAGCTTAAGCTTTT 752
Db 734 CCCGCTCTATGAAGAACGGATCTTAACCTTTT 767

RESULT 8
BI772626 812 bp mRNA linear EST 25-SEP-2001
LOCUS 603060879F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5210299 5',
DEFINITION mRNA sequence.
ACCESSION BI772626
VERSION BI772626
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE NIH-MGC http://mhc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabps-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11527 row: m column: 20
High quality sequence start: 2
High quality sequence stop: 808.
FEATURES
source Location/Qualifiers
1. 812
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5210299"
/lab_host="PH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site.1: NotI; Site.2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT 217 a 186 c 221 g 188 t
ORIGIN

Query Match 29.1%; Score 713.2; DB 13; Length 812;
Best Local Similarity 91.0%; Pred. No. 2.7e-179;
Matches 801; Conservative 0; Mismatches 3; Indels 76; Gaps 1;

QY 1 TGGAGTGGGAGCTCAAGCAGGATCTTCCGAGTCCCTGGCATCTCAGAACCTTCAACT 60
Db 9 TGGAGTGGGAGCTCAAGCAGGATCTTCCGAGTCCCTGGCATCTCAGAACCTTCAACT 68
```

```
QY 61 CTGGAGGCAATGGGTGGAAGAAAGATGACTGCTGAGTTCCTGGAAGAAACAGACACC 120
Db 69 CTGGAGGCAATGGGTGGAAGAAAGATGACTGCTGAGTTCCTGGAAGAAACAGACACC 128
QY 121 AACATCCGGAAGAAACCTTCATTTTATGGAAGTCTGGATCAGGAGCTTTCTCAGAAGTT 180
Db 129 AACATCCGGAAGAAACCTTCATTTTATGGAAGTCTGGATCAGGAGCTTTCTCAGAAGTT 188
QY 181 TTCTGTGTGAAGAAAGAACTGACTGGGAAGCTTTTCTCTGGAAGTCAATCAAGAACTCA 240
Db 189 TTCTGTGTGAAGAAAGAACTGACTGGGAAGCTTTTCTCTGGAAGTCAATCAAGAACTCA 248
QY 241 CCTGCTTCCGGGACAGCAGCCTGGAGAATGAGATTGCTGTGTTGAAAAGATCAAGCAT 300
Db 249 CCTGCTTCCGGGACAGCAGCCTGGAGAATGAGATTGCTGTGTTGAAAAGATCAAGCAT 308
QY 301 GAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCACCTACTACCTGGTCTATG 360
Db 309 GAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCACCTACTACTGCTGCTATG 368
QY 361 CAGCTTTGTTCTGTGGGAGCTCTTTGACCGGATCTCGGAGCGGGGTGTCTACACAGAG 420
Db 369 CAGCTTTGTTCTGTGGGAGCTCTTTGACCGGATCTCGGAGCGGGGTGTCTACACAGAG 428
QY 421 AAGGATGCCAGTCTGGTGATCCAGCAGGTCTTGTGCGCAGTGAATACCTACATGAAAT 480
Db 429 AAGGATGCCAGTCTGGTGATCCAGCAGGTCTTGTGCGCAGTGAATACCTACATGAAAT 488
QY 481 GGCATCGTCCACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC 540
Db 489 GGCATCGTCCACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC 548
QY 541 TCTAAGATCATGATGACTGTTGGTCTGTCCAAAGATGGAACAGATGGCATCATGTCC 600
Db 549 TCTAAGATCATGATGACTGTTGGTCTGTCCAAAGATGGAACAGATGGCATCATGTCC 608
QY 601 ACTGCCGTGGAGCCCGAGCTAGTGTGCTCCAGAGTGTGCGCCAGAACCCCTACAGC 660
Db 609 ACTGCCGTGGAGCCCGAGCTAGTGTGCTCCAGAGTGTGCGCCAGAACCCCTACAGC 636
QY 661 AAGGCTGTGGATTGCTGTGCTCCATCGGCTCATCACCTACATATTTGCTCTGTGATACCCC 720
Db 637 -----GCTCTGTGGATACCCC 652
QY 721 CCGTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAAGATCAAGAGGGCTACTATGAG 780
Db 653 CCATTCATGAAGAAACGGAGTCTAAGCTTTTCGAGAAAGATCAAGAGGGCTACTATGAG 712
QY 781 TTTGAGTCTCCATTCTGGGATGACATTTCTGAGTCAGCAGGAGCTTTATTTGGCAGCTTG 840
Db 713 TATGAGTCTCCATTCTGGGATGACATTTCTGAGTCAGCAGGAGCTTTATTTGCCACTTG 772
QY 841 CTTGAGAAGGATCGACAGCAGCGGTACACCTGTGAGAAGG 880
Db 773 CTTGAGAAGGATCGACAGCAGCGGTACACCTGTGAGAAGG 812

RESULT 9
BI772626 1129 bp mRNA linear EST 05-MAR-2002
LOCUS 6575096 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5732298
DEFINITION 5', mRNA sequence.
ACCESSION BI772626
VERSION BI772626
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE NIH-MGC http://mhc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
```

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue procurement: Invitrogen  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12734 row: c column: 19  
High quality sequence start: 11  
High quality sequence stop: 651.  
Location/Qualifiers  
1. .1129  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5732298"  
/tissue\_lib="NIH\_MGC\_124"  
/tissue\_type="hippocampus"  
/lab\_host="DH10B"  
/note="organ: brain; Vector: pCMV-SPORT6; Site\_1: EcorV  
(destroyed); Site\_2: NotI; RNA source male hippocampus,  
age 27. Library is oligo-dT primed and directionally  
cloned (EcorV site is destroyed upon cloning). Average  
insert size 1.4 kb, insert size range 0.9-4 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 012."  
287 a 308 c 294 g 240 t

BASE COUNT  
ORIGIN

Query Match 29.0%; Score 710.2; DB 14; Length 1129;  
Best Local Similarity 95.1%; Pred. No. 2.1e-178;  
Matches 800; Conservative 0; Mismatches 28; Indels 13; Gaps 6;

```
QY 1 TGGAGTGGAGCTCAAGCAGGATTCCTCCGAGTCCCTGGCATCTCAGAACTTCAACT 60
|||||
DB 36 TGGAGTGGAGCTCAAGCAGGATTCCTCCGAGTCCCTGGCATCTCAGAACTTCAACT 95
|||||
QY 61 CTGGAGCCTGGTGGAGGAGAGATGACTGCAGTCTCTGGAAGAAACAGACCACC 120
|||||
DB 96 CTGGAGCCTGGTGGAGGAGAGATGACTGCAGTCTCTGGAAGAAACAGACCACC 155
|||||
QY 121 AACATCCGGAAACCTTCATTTTATGGAAGTCTGGATCAGGAGCTTCTCAGAAAGTT 180
|||||
DB 156 AACATCCGGAAACCTTCATTTTATGGAAGTCTGGATCAGGAGCTTCTCAGAAAGTT 215
|||||
QY 181 TTCTCTGGTGAAGCAAGACGACTGCTGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA 240
|||||
DB 216 TTCTCTGGTGAAGCAAGACGACTGCTGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA 275
|||||
QY 241 CTGTCCTCTCCGGACAGCCTGGAGATGAGATGCTGCTTGTGTTGAAAAGATCAAGCAT 300
|||||
DB 276 CTGTCCTCTCCGGACAGCCTGGAGATGAGATGCTGCTTGTGTTGAAAAGATCAAGCAT 335
|||||
QY 301 GAAAACATTTGTACCTCGGAGGACATCTATGAGAGCACCACCCTACTTACCTGGTCATG 360
|||||
DB 336 GAAAACATTTGTACCTCGGAGGACATCTATGAGAGCACCACCCTACTTACCTGGTCATG 395
|||||
QY 361 CAGCTTTTCTGGTGGGAGCTCTTTGACCGGATCTCTGGAGCGGGGTCTCTACACAGAG 420
|||||
DB 396 CAGCTTTTCTGGTGGGAGCTCTTTGACCGGATCTCTGGAGCGGGGTCTCTACACAGAG 455
|||||
QY 421 AAGGATCCAGTCTGTGTATCCAGCAGCTTTGTCGGCAGTCAAAATACCTACATCAGAAAT 480
|||||
DB 456 AAGGATCCAGTCTGTGTATCCAGCAGCTTTGTCGGCAGTCAAAATACCTACATCAGAAAT 515
|||||
QY 481 GGCATCGTCCACAGAGACTTAAAGCCGCAAAACCTTGTACCTTTACCTTACCCCTGAAGAGAAC 540
|||||
DB 516 GGCATCGTCCACAGAGACTTAAAGCCGCAAAACCTTGTACCTTTACCTTACCCCTGAAGAGAAC 575
|||||
QY 541 TCTAAGATCATGATCACTGACTTGGTCTGTCCAGATGGAGACAGATGGCATCATGTCC 600
```

```
Db 576 TCTAAGATCATGATCACTGACTTTGGTCTGTCCAAGATGGAACAGATGGCATCATGTCC 635
|||||
QY 601 ACTGCCCTGTGGACCCAGGCTAGCTGGCTCCAGAACTGTGGCCAGAAACCTACAGC 660
|||||
DB 636 ACTGCCCTGTGGACCCAGGCTAGCTGGCTCCAGAACTGTGGCCAGAAACCTACAGC 695
|||||
QY 661 AAGGC---TGTGGATTTGCTGGTCCATCGGGCTCATCA-CCTACATATTTGCTCTGTGGATA 716
|||||
DB 696 AAGGCCTGTGGATTTGCTGGTCCATCGGGGCTATCACCCTACCTATTTGCTCTGTGGATA 755
|||||
QY 717 ---CCCCCGCTTCTATGAA-AAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCT 772
|||||
DB 756 ACCCCCCCGCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCT 815
|||||
QY 773 ACTATGAG---TTTGAGTCTCCATTTCTGGG--ATGACATTTCTGAGTCAGCCAAAGGACTT 827
|||||
DB 816 TAACATGAGTTTGTAGTCTCCATTCCTGGGATGAACATTCCTGAGTCAGCCCGGAGCT 875
|||||
QY 828 T 828
|
DB 876 T 876
```

RESULT 10  
BG715920  
LOCUS  
DEFINITION 602676667F1 NIH\_MGC\_96 Homo sapiens cDNA clone IMAGE:4799224 5',  
mRNA sequence.  
ACCESSION BG715920  
VERSION BG715920.1 GI:13995107  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 824)  
AUTHORS NTH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM10688 row: e column: 17  
High quality sequence stop: 766.

FEATURES  
source

Location/Qualifiers  
1. .824  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4799224"  
/tissue\_lib="NIH\_MGC\_96"  
/tissue\_type="hypothalamus"  
/lab\_host="DH10B"  
/note="organ: brain; Vector: pBluescriptR (modified  
pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcgag  
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',  
size-selected for average insert size 2.3 kb and  
normalized to ROT 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NHGRI), National  
Institutes of Health). Note: this is a NIH\_MGC Library."  
223 a 194 c 219 g 188 t

Query Match

28.7%; Score 701.4; DB 12; Length 824;

Best Local Similarity 96.9%; Pred. No. 4e-176; Matches 747; Conservative 0; Mismatches 21; Indels 3; Gaps 3;

QY	1	TGGAGTGGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCTCTCAGAAGCTTCAACT	60
Db	38	TGGAGTGGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCTCTCAGAAGCTTCAACT	97
QY	61	CTGAGGCAATGGTCCGAAGGAGAGATGACTGCTGGATCAGGAGCTTCTCAGAAGTT	120
Db	98	CTGAGGCAATGGTCCGAAGGAGAGATGACTGCTGGATCAGGAGCTTCTCAGAAGTT	157
QY	121	AACATCCGAAAAACCTTCATTTTATGGAAGTCTCTGGATCAGGAGCTTCTCAGAAGTT	180
Db	158	AACATCCGAAAAACCTTCATTTTATGGAAGTCTCTGGATCAGGAGCTTCTCAGAAGTT	217
QY	181	TTCTGGTGAAGCAAGACTGACTGGGAAGCTTTTGGCTCTCTGAAAGTCAAGAGTCA	240
Db	218	TTCTGGTGAAGCAAGACTGACTGGGAAGCTTTTGGCTCTCTGAAAGTCAAGAGTCA	277
QY	241	CCTGCCTTCCGGGACAGCCTGGAGATGAGATGCTGTGTTGAAAGATCAAGCAT	300
Db	278	CCTGCCTTCCGGGACAGCCTGGAGATGAGATGCTGTGTTGAAAGATCAAGCAT	337
QY	301	GAACAATGTGACCTGGAGGACATCTATGAGAGCACCACCTACTTACCTGTGTCATG	360
Db	338	GAACAATGTGACCTGGAGGACATCTATGAGAGCACCACCTACTTACCTGTGTCATG	397
QY	361	CAGCTGTCTTCTGGTGGGAGCTCTTTGACCGGATCCTGGAGGGGTCTACACAGAG	420
Db	398	CAGCTGTCTTCTGGTGGGAGCTCTTTGACCGGATCCTGGAGGGGTCTACACAGAG	457
QY	421	AAGGATGCCAGTCTGGTGTACAGCAGGCTTTGTGGCAGTGAATACCTACATGAGAT	480
Db	458	AAGGATGCCAGTCTGGTGTACAGCAGGCTTTGTGGCAGTGAATACCTACATGAGAT	517
QY	481	GGCATCGTCCACAGACACTTAAAGCCCGAAAAACCTTACCTTACCCCTGAAGAGAC	540
Db	518	GGCATCGTCCACAGACACTTAAAGCCCGAAAAACCTTACCTTACCCCTGAAGAGAC	577
QY	541	TCTAAGATCATGATCACTGACTTTGGTCTGTCAGAGTGGACAGAAATGCGATATGTC	600
Db	578	TCTAAGATCATGATCACTGACTTTGGTCTGTCAGAGTGGACAGAAATGCGATATGTC	637
QY	601	ACTGCCTGTGGGACCCAGGCTACG-TGGCTCCAGAGTGTGGCCAG-AAACCTTACA	658
Db	638	ACTGCCTGTGGGACCCAGGCTACGTTGGCTCCAGAGTGTGGCCAGAAACCTTACA	697
QY	659	GCAAGCTGTGGATGCTGGTCCATCGGCGTCACTACCTTACATATTTGCTGTGGATACC	718
Db	698	GCAAGCTGTGGATGCTGGTCCATCGGCGT-ATCACCTTAACATATGCTCTGTGGATACC	756
QY	719	CCCGCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAGATCAAGGAGG	769
Db	757	CCCATCTCTATGAAGAAACGGAGTCTTACGCTTTTCGAAAAATCAAGAGGG	807

RESULT 11  
BI753035  
LOCUS 603025844Fl NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5196323 5',  
DEFINITION mRNA sequence.  
ACCESSION BI753035  
VERSION BI753035.1 GI:15744613  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 731)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.

Email: [ccapbs-r@mail.nih.gov](mailto:ccapbs-r@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LAM11491 row: g column: 12  
High quality sequence stop: 729.

FEATURES  
source  
1..731  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5196323"  
/clone\_lib="NIH\_MGC\_114"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH\_MGC Library."

BASE COUNT 190 a 182 c 191 g 168 t  
ORIGIN

Query Match 28.6%; Score 699; DB 13; Length 731;  
Best Local Similarity 99.7%; Pred. No. 1.6e-175;  
Matches 721; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY	1	TGGAGTGGGAGCTCAAG-CAGGATTTCTTCCCGAGTCCCTGGCATCTCTCAGAAGCTTCAAC	59
Db	1	TGGAGTGGGAGCTCAAGCCAGGATTTCTCCCGAGTCCCTGGCATCTCTCAGAAGCTTCAAC	60
QY	60	TCTGGAGCAATGGTGGAAAGGAGAGATGACTGAGTTCCTGGAAAGAACAGACAC	119
Db	61	TCTGGAGCAAT-GGTGGAAGGAGAGATGACTGAGTTCCTGGAAAGAACAGACAC	119
QY	120	CAACATCCGAAAAACCTTCATTTTATGGAAGTGTGGGATCAGGAGCTTCTCAGAAGT	179
Db	120	CAACATCCGAAAAACCTTCATTTTATGGAAGTGTGGGATCAGGAGCTTCTCAGAAGT	179
QY	180	TTTCTGTGTGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAAGTGCATCAAGAAGTC	239
Db	180	TTTCTGTGTGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAAGTGCATCAAGAAGTC	239
QY	240	ACCTGCCCTTCCGGGACAGCAGCTGGGAATGAGATGCTGTGTTGAAAGATCAAGCA	299
Db	240	ACCTGCCCTTCCGGGACAGCAGCTGGGAATGAGATGCTGTGTTGAAAGATCAAGCA	299
QY	300	TGAAACATTTGACCCCTGGAGGACATCTATGAGAGCACCACCTACTACCTGGTGCAT	359
Db	300	TGAAACATTTGACCCCTGGAGGACATCTATGAGAGCACCACCTACTACCTGGTGCAT	359
QY	360	GCAGCTGTCTGTGGTGGGAGCTTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGA	419
Db	360	GCAGCTGTCTGTGGTGGGAGCTTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGA	419
QY	420	GAAGGATGCCAGTCTGGTGTATCCAGAGGCTTCTTCCGCGAGTGAATACCTACATGAGAA	479
Db	420	GAAGGATGCCAGTCTGGTGTATCCAGAGGCTTCTTCCGCGAGTGAATACCTACATGAGAA	479
QY	480	TGGCATCGTCCACAGAGACTTTAAAGCCCGAAAAACCTTTACCTTACCCCTGAAGAGAA	539
Db	480	TGGCATCGTCCACAGAGACTTTAAAGCCCGAAAAACCTTTACCTTACCCCTGAAGAGAA	539
QY	540	CTCTAAGATCATGATCACTGACTTTGGTGTCTGTCGAAGTGAACAGAAATGCGATCATGTC	599
Db	540	CTCTAAGATCATGATCACTGACTTTGGTGTCTGTCGAAGTGAACAGAAATGCGATCATGTC	599

```
QY 600 CACTGCTGTGGGACCCAGGCTAGTGGCTCCAGAAAGTGTGGCCCAAGAAACCCCTACAG 659
|||||
Db 600 CACTGCTGTGGGACCCAGGCTAGTGGCTCCAGAAAGTGTGGCCCAAGAAACCCCTACAG 659
|||||
QY 660 CAAGGCTGTGGATTCTGTGTCATCGGCGTCAATCACTACACCTACATATTGCTCTGGATACCC 719
|||||
Db 660 CAAGGCTGTGGATTCTGTGTCATCGGCGTCAATCACTACACCTACATATTGCTCTGGATACCC 719
|||||
QY 720 CCC 722
|||
Db 720 CCC 722

RESULT 12
BQ934044
LOCUS AGENCOURT_8802332 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6336727
DEFINITION 5', mRNA sequence.
ACCESSION BQ934044
VERSION BQ934044.1 GI:22349427
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 905)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.
cDNA Library Preparation: ResGen, Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM13801 row: d column: 08
High quality sequence stop: 689.
Location/Qualifiers
1..905
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6336727"
/clone.lib="NIH_MGC_130"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Otcysts; Vector: pCMV-SPORT6.1.cdb; Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."
BASE COUNT 238 a 217 c 240 g 210 t
ORIGIN

Query Match 26.7%; Score 653.6; DB 14; Length 905;
Best Local Similarity 90.0%; Pred. NO. 2.5e-163;
Matches 745; Conservative 0; Mismatches 79; Indels 4; Gaps 4;

QY 31 GAGTCCCTGGCATCTCAGAAAGTCTCACTCTGGAGCAATGGGTGGAAGGAAGCAAGAT 90
|||||
Db 68 GAGCCCTCGGCTCTCAGGCGACTTCACTCTGGAGGACATGGGCGGTAAGGAGGAG 127
|||||
QY 91 GACTGCAGTTCCTGGAAGAAACAGACCAACATCCGGAACACCTTCATTTTATGAA 150
|||||
Db 128 GACTGCAGTTCCTGGAAGAAACAGACCAACATCCAGGAACCTTCATCTTCATGAA 187
|||||
QY 151 GTGCTGGGATCAGGAGCTTTCTCAGAAGTTTCTCGTGGTGAAGCAAGACTGACTGGGA 210
|||||
Db 188 GTGCTGGGATCAGGAGCTTTCTCAGAAGTTTCTCGTGGTGAAGCAAGACTGACTGGGA 247
|||||
QY 211 CTCTTTGCTCTGAAGTGCATCAAGAAGTCACTCTCCGCGACAGCAGCGCTGGAGAT 270
|||||
```

```
Db 248 CTCTTTGCTCTGAATGTATCAAGAAGTCAACAGCCTTCGGGACAGCAGCCCTAGAGAT 307
|||||
QY 271 GAGATCTGCTGTGTGAAAGATCAAGCATGAAACATTTGTGACCCCTGGAGGACATCTAT 330
|||||
Db 308 GAGATCGCTGTGTGAAAGATCAAGCATGAAACATTTGTGACCCCTGGAGGACATCTAT 367
|||||
QY 331 GAGAGCACCAACCACTACTACCTGGTGCATGCAGCTGTTTCTGGTGGGAGCTCTTTGAC 390
|||||
Db 368 GAGAGCACCAACCACTACTACCTGGTGCATGCAGCTGTTTCTGGAGTGAGCTTTTGAC 427
|||||
QY 391 CGGATCCTGGAGCGGGGTGTCTACAGAGAAAGGATGCCAGTCTGGTGATCCAGCAGGTC 450
|||||
Db 428 CGGATCCTAGAGCGGTGTCTACAGAGAAAGGATGCCAGTCTGGTGATCCAGCAGGTC 487
|||||
QY 451 TTGTCGGCAGTGAATACCTACATGAGAATGGCATCTCCACAGAGACTTTAAAGCCCGAA 510
|||||
Db 488 TTGTCGGCAGTGAATACCTTCATGAGAATGGCATCTCCACAGAGATCTAAAGCCCTGAA 547
|||||
QY 511 AACCTGCTTACCTTACCCCTGAAGAACTCTAAGATCATGATCACTGACTTTGGTCTG 570
|||||
Db 548 AACCTGCTTACCTTACCCCTGAAGAACTCTAAGATCATGATCACTGACTTTGGTCTA 607
|||||
QY 571 TCCAGATGGAACAGATGCGATCATGTCACCTGCTGCGACCCAGCAGCTAGCTGGCT 630
|||||
Db 608 TCCAGATGGAAGAGATGAGTGTGTCACAGCTTGTGGACCCAGCAGCTAGCTGGCT 667
|||||
QY 631 CCAGAGTGTGTGGCCCAAGAAACCTACAGCAAGGCTGTGGATTGCTGCTCCATCGCGCT 690
|||||
Db 668 CCAGAGTGTGTGGCCCAAGAAACCTACAGTAAGCTGTGCTGCTGCTGCTGCTGCTG 727
|||||
QY 691 ATCACTACATATTTGCTGCTGATACCCCGCTTCTATGAAGAAAGGAGTCTAAGC-T 749
|||||
Db 728 ATCACTACATATTTGCTGCTGATACCCCGCTTCTATGAAGAAAGGAGTCTAAGC-T 787
|||||
QY 750 TTTCGAGAAGATCAAGGAGGCTACTATGA-GTTTGAAGTCTCCATTTCTGGATGACATTT 808
|||||
Db 788 TTTTGAAGATCAAGAGAGGCTACTAGAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTTT 847
|||||
QY 809 CTGAGTCA-GCCAGGACTTTATTTGCCA-CTTGCTTTGAGAAGATCC 854
|||||
Db 848 CTGAGTCAAGGAGGATTTATTTGCCATCTGCTGGAGAAAGGACCC 895
|||||

RESULT 13
BQ293660
LOCUS 602390529F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4502479 5',
DEFINITION mRNA sequence.
ACCESSION BQ293660
VERSION BQ293660.1 GI:13053536
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 818)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM10371 row: i column: 08
High quality sequence stop: 765.
Location/Qualifiers
1..818

FEATURES
source
```

RESULT 14  
BT667965

QY 360 GCAGCTGTTCTGGTGGGAGCTCTTTGACCGSATCC--TGGAGCGGGTGTCTACACA 417  
Db 396 GCAGCTGTTCTGGTGGGAGCTCTTTGACCGGATCATGGCAGCGGGTGTCTACACA 455  
QY 418 GAGAAAGATGCCAGTCTGGTGATCCAGCAGGTCTTGTGGCAGTGAATACCTACATGAG 477  
Db 456 GAGAAAGATGCCAGTCTGGTGATCCAGCAGGTCTTGTGGCAGTGAATACCTACATGAG 515  
QY 478 AATGGCATGTCACAGAGACTTAAAGCCG-AAAACCTGCTTTACCTTACCCCTGAAGA 536  
Db 516 AATGGCATGTCACAGAGACTTAAAGCCGGAAGAAACCTGCTTTACCTTACCCCTGAAGA 575  
QY 537 GAATCTTAAGATCATGATCACTGACTTGTGCTGCTCAAGATGGAACAGATGGCATC-A 595  
Db 576 GAATCTTAGGATCATGATCACTGACTGTGCTGCTCAAGATGGAACAGATGGCATCAA 635  
QY 596 TGTCCACTGCTGTGGGACCC 616  
Db 636 TGTCCACTGCTGTGGGACCC 656

RESULT 15  
BI834635 554 bp mRNA linear EST 04-OCT-2001  
DEFINITION 603090418F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5229245 5',  
mRNA sequence.  
ACCESSION BI834635  
VERSION BI834635.1 GI:15946185  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 554)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
DNA distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L14M11577 row: c column: 06  
High quality sequence stop: 554.  
Location/Qualifiers  
1..554  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5229245"  
/lab\_host="NIH\_MGC\_120"  
/lab\_host="DH10B"  
/note="Organ: pooled pancreas and spleen; Vector:  
pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
source anonymous pool of spleen and pancreas from 28 yo  
male. Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 1.5 kb, insert size range 1-2.5 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 025. Note: this is a NIH\_MGC Library."

BASE COUNT 156 a 128 c 143 g 127 t  
ORIGIN

Query Match 21.3%; Score 522; DB 13; Length 554;  
Best Local Similarity 98.0%; Pred. No. 3.1e-128;  
Matches 543; Conservative 0; Mismatches 0; Indels 11; Gaps 1;

QY 82 GAAGAAGATGACTGCTGCTGGAGAAACAGACCACCAATCCGGAACACCTTCATT 141

Db 1 GAAGAAGATGACTGCTGCTGCTGGAAGAAACAGACCACCAACATCCGGAACACCTTCATT 60  
QY 142 TTTATGGAAGTGTGGGATCAGGAGCTTTCTCAGAGTCTTTCTGTGTGAAGCAAGACTG 201  
Db 61 TTTATGGAAGTGTGGGATCAGGAGCTTTCTCAGAGTCTTTCTGTGTGAAGCAAGACTG 120  
QY 202 ACTGGGAAGCTCTTTGCTGCTGAAAGTGCATCAAGAAAGTCACTGCTCTCCGCGGACACGACG 261  
Db 121 ACTGGGAAGCTCTTTGCTGCTGAAAGTGCATCAAGAAAGTCACTGCTCTCCGCGGACACGACG 180  
QY 262 CTGGAAGATGAGATGTGTGTTGAAAGATCAAGCATGAAACATTTGTGACCCCTGGAG 321  
Db 181 CTGGAAGATGAGATGTGTGTTGAAAGATCAAGCATGAAACATTTGTGACCCCTGGAG 240  
QY 322 GACATCTATGAGAGCACACCCACTACTACCTGCTGTCATGCAGCT-----TGTGTT 370  
Db 241 GACATCTATGAGAGCACACCCACTACTACCTGCTGTCATGCAGCTGTAAGTAAAGTGTGTT 300  
QY 371 CTGTTGGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGGATGCCA 430  
Db 301 CTGTTGGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGGATGCCA 360  
QY 431 GTCGTGTGATCCAGCAGGCTCTTGTGCGCAGTGAATACCTACATGAGAAATGGCATGCTCC 490  
Db 361 GTCGTGTGATCCAGCAGGCTCTTGTGCGCAGTGAATACCTACATGAGAAATGGCATGCTCC 420  
QY 491 ACAGAGACTTTAAAGCCCGAAACCTGCTTTACCTTTACCCCTGAAGAGAACTCTAAGATCA 550  
Db 421 ACAGAGACTTTAAAGCCCGAAACCTGCTTTACCTTTACCCCTGAAGAGAACTCTAAGATCA 480  
QY 551 TGATCACTGACTTTGGTCTGTCCAAGATGGAACAGAAATGGCATGCTCCACTGCTGTG 610  
Db 481 TGATCACTGACTTTGGTCTGTCCAAGATGGAACAGAAATGGCATGCTCCACTGCTGTG 540  
QY 611 GGACCCCGAGGCTAC 624  
Db 541 GGACCCCGAGGCTAC 554

Search completed: March 14, 2003, 15:22:05  
Job time : 3891 secs